STOCHASTIC SEARCH EQUILIBRIUM

By

Giuseppe Moscarini and Fabien Postel-Vinay

February 2010

COWLES FOUNDATION DISCUSSION PAPER NO. 1754

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

Stochastic Search Equilibrium*

Giuseppe Moscarini†
Yale University
and
NBER

Fabien Postel-Vinay‡
University of Bristol
and
Paris School of Economics

February 2010

Abstract

We study a stochastic economy where both employed and unemployed workers search randomly for labor contracts posted by firms, while aggregate productivity is subject to persistent shocks. Our exercise provides the first dynamic stochastic general equilibrium analysis of a popular class of search wage-posting models, drawing in part from the literature on recursive contracts under moral hazard. Each firm offers and commits to a (Markov) contract, which specifies a wage contingent on all payoff-relevant states, but must pay equally all of its workers, who have limited commitment and are free to quit at any time. An equilibrium of this contract-posting game is Rank-Preserving [RP] if larger firms offer a larger value to their workers in all states of the world. We find two sufficient (but not necessary) conditions for every equilibrium to be RP: either firms only differ in their initial size, or they also differ in their fixed idiosyncratic productivity but more productive firms are initially weakly larger, in which case turnover is always efficient, as workers always move from less to more productive firms. In both cases, the ranking of firm sizes never changes on the RP equilibrium path, a property that has three useful implications. First, the stochastic dynamics of firm size, uniquely pinned down in equilibrium, provide an intuitive explanation for the empirical finding that large employers are more cyclically sensitive (Moscarini and Postel-Vinay, 2009). Second, contracts are unique in RP equilibrium. Third, RP equilibrium computation is tractable, and we construct and simulate calibrated examples.

Keywords: Equilibrium Job Search, Dynamic Contracts, Stochastic Dynamics.
JEL codes: J64, J31, D86.

---

*Earlier versions of this paper circulated under the title "Non-stationary search equilibrium". We acknowledge useful comments to earlier drafts of this paper from seminar and conference audiences at numerous venues. We also wish to thank Ken Burdett, Dale Mortensen and Rob Shimer for constructive discussions of earlier versions of this paper. The usual disclaimer applies.

†Address: Department of Economics, Yale University, PO Box 208268, New Haven CT 06520-8268. Tel. +1-203-432-3596. E-mail `giuseppe.moscarini@yale.edu`. Web `http://www.econ.yale.edu/faculty1/moscarini.htm`

‡Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK. Tel: +44 117 928 8431. E-mail `Fabien.Postel-Vinay@bristol.ac.uk`. Web `www.efm.bris.ac.uk/economics/staff/vinay/`. Postel-Vinay is also affiliated with CEPR (London) and IZA (Bonn).

1 Introduction

We study the equilibrium dynamics of a frictional labor market where firms offer and commit to employment contracts and workers search randomly on and off the job for those contracts, while aggregate productivity is subject to persistent shocks. In a broad sense, our exercise sheds light on the long-term contracts that emerge in a market equilibrium environment, in the presence of moral hazard and aggregate uncertainty. More specifically, we perform the first analysis of aggregate stochastic dynamics in a popular class of search wage-posting models, originating with Burdett and Mortensen (1998, henceforth BM). By providing a coherent formalization of the hypothesis that cross-sectional wage dispersion is largely a consequence of labor market frictions, the BM model has started a fruitful line of research in the analysis of wage inequality and worker turnover, as the vibrant and empirically very successful literature organized around that hypothesis continues to show (see Mortensen, 2003 for an overview).

That literature, however, is invariably cast in deterministic steady state. Ever since the first formulation of the BM model, job search scholars have regarded the characterization of its out-of-steady-state behavior as a daunting problem, essentially because one of the model's state variables, which is also the main object of interest, is the endogenous *distribution* of wage (or job value) offers. This is an infinite-dimensional object, endogenously determined in equilibrium as the distribution across firms of offer strategies that are mutual best responses, and evolving stochastically with the aggregate impulse. In this paper, we solve this problem.

We show that, under mild sufficient conditions, the economy under investigation has a unique equilibrium. In this equilibrium, the workers' ranking of firms is the same in all aggregate states — what we call a *Rank-Preserving Equilibrium* (RPE). The sufficient conditions are that firms either are equally productive, or differ in the permanent component of their productivity and the more productive they are, the (weakly) more workers they initially employ — for example, they all start empty. In the latter case, in RPE more productive firms offer a larger value and employ more workers at all points in time: when given a chance, a worker always moves from a less into a more productive firm, so that the equilibrium allocation of employment is constrained efficient. This parallels a similar property of BM's static equilibrium.

To illustrate and qualify our contributions we now provide details about the economy that we study, our solution method, and the nature of the unique equilibrium. Infinitely lived and risk neutral firms and workers come in contact infrequently. Firms produce homogenous output with labor in a linear technology, which may permanently differ across firms. Aggregate multiplicative TFP shocks affect labor productivity as well as the job contact rates, on and off the job, the exogenous job destruction rate, and the value of leisure.

In this economy, the constrained efficient allocation is easily characterized. When given the opportunity, an employed worker is moved from a less productive to a weakly more productive firm. We abstract from issues of entry and exit and restrict attention to parameter configurations such that, as in BM, employment is always preferred to unemployment. This efficient turnover behavior describes a simple Markov process for the evolution of the firm size distribution. If we shut down aggregate shocks, from any initial condition which gives rise to RPE this process converges deterministically to the size distribution that BM found by solving directly for the stationary distribution. In a stochastic environment, for any history of aggregate shocks, we can solve in closed form for the path of the socially optimal distribution of employment across firms, thus of the size of each firm type.

The next step is equilibrium analysis. We assume that firms offer and commit to a contract which conditions the wage on all possible relevant states and is subject to an *equal-treatment constraint*: it must pay the same wage in a given period to all of its employees, whether incumbent, newly hired from unemployment or from employment. This constraint indeed defines the boundaries of a firm. Workers cannot commit not to quit to other jobs when the opportunity arises, or to unemployment whenever they please, so commitment is one-sided and firms face a standard moral hazard problem. Contract offers are privately observed only by the recipients, thus deviations cannot be detected by other players.

We look for a Sequential Nash equilibrium of this contract-posting game. We find the largest state space on which equilibrium contracts can be conditioned. For tractability, we then restrict attention to a Markov Perfect equilibrium, where wages depend only on payoff-relevant states: two exogenous, the productivity of the firm and the state of aggregate productivity, one endogenous to the firm, its current size, and one endogenous to the economy but exogenous to the firm, the distribution of employment across all firms. A firm must track this infinitely-dimensional object in order to know the distributions of competing offers and of values earned by currently employed workers, thus how much recruitment and retention its own contract will generate. Equilibrium only imposes two very weak restrictions on these two distributions: they must have no atoms and a connected support. Yet we are able to establish that at most one Markov perfect equilibrium exists, characterize that unique equilibrium and show that it decentralizes the constrained efficient allocation.

The key step in our analysis is a comparative dynamics property of the best-response. We show that, at any node in the game and for any distribution of offers made by other firms and of values earned by employed workers, a more productive and/or larger firm optimally offers a contract that pays its existing and new workers a larger value. Therefore, if firms are homogeneous, or if more productive firms are initially no smaller, then no firm wants to break ranks in the distribution of competing offers, which then coincides with the given distribution of firm productivities or initial sizes. This immediately implies our main result that equilibrium, if it exists, is unique and RP, thus constrained efficient.

The intuition behind this comparative dynamics property of the best-response contract parallels a single-crossing property of the static BM model. There, a more productive firm gains more from employing a worker, hence wants to (and can) pay a higher wage. In addition, under the equal treatment constraint, the effect of the wage on retention is proportional to own size, while that on hiring is independent of size. Finally, size increases in the wage due to its effect on recruitment and retention. Thus, ceteris paribus, a larger firm also wants to pay more. This intuition does not extend immediately to our dynamic stochastic setting, because firm size is an evolving state variable with a given initial condition. The dynamic incentives of a firm to pay more or less T periods down the road depend on its size at T, which in turn is determined by the contract given the history of shocks before time T. If more productive firms are initially weakly larger, this initial size ranking self-perpetuates on the equilibrium path, as required by RPE. More productive firms always pay and employ more. While this outcome is intuitive and natural, we find it remarkable that it is unique despite the strong strategic complementarity of a wage-posting game.

As a by-product of this analysis we offer a methodological contribution, namely the first (to the best of our knowledge) theory of Monotone Comparative Dynamics in a dynamic stochastic decision problem. In our setting, firms solve a fully dynamic problem in a changing environment. In the sequential formulation of this problem, the choice set is an infinite sequence (a stochastic process), a case that the theory of Monotone Comparative Statics (Topkis, 1998) does not cover. The objective function of the one-step Bellman maximization contains the value function of the problem, whose properties are ex ante unknown. We establish that the Bellman operator of the contract-posting problem is a contraction on the space of functions that satisfy some single-crossing and convexity properties, which are then inherited by the value function. We can then apply Monotone Comparative Statics to the Bellman equation and a forward induction argument to prove the Monotone Comparative Dynamics properties of the best response illustrated above. The same logic can be applied in many other settings: for example, one may ask in a stochastic growth model whether the socially optimal level of investment is decreasing at all states and dates in the initial stock of capital, in the labor share, or in risk aversion; in each case, one must track the effect of the parameter change on the endogenous state variable (capital) along the entire optimal path.

One final, and crucial, benefit of RPE is its tractability and computability. As employment allocation is constrained efficient, we can solve for it analytically, which affords a direct proof of uniqueness of the equilibrium allocation. To compute equilibrium contracts, we further establish that the firm value function and best-response contract are differentiable in firm productivity and size, a non-trivial property because the distributions of offers that the firm responds to may lack a density at many points. Next, using first-order conditions and Euler equations solved by optimal contracts, we obtain our third and final contribution: we establish uniqueness and a sufficient condition for existence of the equilibrium contract, further suggesting an algorithm to compute state-contingent wages, thus a constructive proof of equilibrium existence. We finally calibrate and simulate the model. This numerical exercise, albeit mostly illustrative, reveals the quantitative potential of the model.

Besides being of intrinsic theoretical interest, our characterization of the dynamics of the BM model opens the analysis of aggregate labor market dynamics as a whole potential new field of application of search/wage-posting models. Unlike the typical representative-agent model, our stochastic version of these models makes predictions about the business cycle behavior of wage distributions, firm size distributions, or patterns of labor reallocation across firms, that we can confront with new empirical evidence from various (often new) data sets presented in Moscarini and Postel-Vinay (2009, MPV09). Our model already explains many of these new facts, most notably that large employers have more cyclical employment, and the results of our numerical analysis suggest extensions to further improve its quantitative performance. More generally, we hope to contribute to a synthesis between the BM contract-posting approach and the "other", equally successful side of the search literature, organized around the matching framework (Pissarides, 2000), initially designed for the understanding of equilibrium unemployment and labor market flows.[1]

The rest of the paper is organized as follows. In Section 2 we lay out the basic environment. In Section 3 we characterize the constrained efficient allocation. In Section 4 we describe and formally define an equilibrium. We then introduce the notion of Rank Preserving Equilibrium in Section 5, where we also state our main result about the generality of RPE and give a characterization of RPE, together with uniqueness and existence results. Section 6 shows simulations of equilibrium paths for wages and employment in a RPE. Finally, Section 7 concludes by discussing future extensions of the model.

[1]Rudanko (2008) and Menzio and Shi (2008) analyze wage contract-posting models with aggregate productivity shocks, where job search is directed. The latter assumption greatly simplifies the analysis by severing the link between the individual firm's contract-posting problem and the distribution of contract offers. This is the main hurdle that we face, and that we resolve by exploiting the emergence of Rank-Preserving Equilibria, while maintaining BM's assumption of random search common to the vast majority of the search literature. From a theoretical viewpoint, we see both programs as fruitful directions of exploration. From a quantitative viewpoint, the directed search approach is focused on the response of the job-finding rate to aggregate shocks, and does not generate a well-defined notion of employer size. Hence, it is silent on the wealth of new evidence that we offer in MPV09 on cyclical patterns of the employer size/growth relationship, and that we envision as central to our understanding of the propagation of aggregate shocks in labor markets.

2 The economy

We study a stochastic economy where firms commit to employment contracts and workers search randomly for those contracts. The special case of a stationary and deterministic economy where contracts are restricted to a constant wage is the BM wage posting model with heterogeneous firm types. We present our model in discrete time, as it affords more clarity in the presentation of the contract posting problem under one-sided commitment as a recursive problem.

The labor market is populated by a unit-mass of workers, who can be either employed or unemployed, and by a unit measure of firms.[2] Workers and firms are risk neutral, infinitely lived, and maximize payoffs discounted with factor $\beta \in (0,1)$. Firms operate constant-return technologies with labor as the only input and with productivity scale $\omega\theta$, where ω is an aggregate component, evolving within some bounded set of values $\Omega \subset \mathbb{R}_+$ according to a discrete-time stationary first-order Markov process $H(d\omega' \mid \omega)$, and θ is a fixed, firm-specific component, distributed across firms θ according to a cdf Γ over $\left[\underline{\theta}, \overline{\theta}\right] \subset R_+$.

The labor market is affected by search frictions in that unemployed workers can only sample job offers sequentially with some probability $\lambda_0^\omega \in (0,1)$ each period. Employed workers earn a wage, are allowed to search on the job, and face a per-period sampling chance of job offers of $\lambda_1^\omega \in (0,1)$. For notational simplicity we will assume uniform sampling of firms by workers, in that any worker receiving a job offer draws the type of the firm from which the offer emanates from the distribution $\Gamma(\cdot)$.[3] All firms of equal productivity θ start out with the same labor force. We denote by $\Lambda_0(\theta)$ the measure of employment initially at firms of productivity at most θ. Each employed worker is separated from his employer and enters unemployment every period with probability $\delta^\omega \in (0,1)$. Note that all these transition probabilities, although exogenous, are allowed to depend on the aggregate state ω.

In each period, the timing is as follows. Given a current state ω of aggregate labor productivity and size (measure of workers employed) L:

1. production and payments take place at all firms in current state ω; the flow benefit b^ω accrues to unemployed workers;

2. the new state ω' of aggregate labor productivity is realized;

[2] We implicitly fix the measure of active firms, thus remaining mostly silent on the question of entry and exit. A simple extension of the model to make it capture entry and exit of firms over the business cycle is illustrated in our companion paper Moscarini and Postel-Vinay (2008, MPV08). Finally, that the mass of firms and workers both have measure one is obviously innocuous and only there to simplify the notation.

[3] When calibrating and simulating the model in Section 6 we allow for non-uniform sampling, in that different types-θ firms have different chances of being sampled by job searchers. This extension is theoretically straightforward and useful in quantitative applications.

3. employed workers can quit to unemployment;

4. jobs are destroyed exogenously with chance $\delta^{\omega'}$;

5. the remaining employed workers receive an outside offer with chance $\lambda_1^{\omega'}$ and decide whether to accept it or to stay with the current employer;

6. each previously unemployed worker receives an offer with probability $\lambda_0^{\omega'}$.

Finally, in order to avert unnecessary complication, we will assume throughout the paper that the distribution of firm types, Γ, has continuous and everywhere strictly positive density over $\left[\underline{\theta}, \overline{\theta}\right]$, and that the initial measure of employment across firm types, Λ_0, is continuously differentiable in θ. Combining those two assumptions, we obtain that the initial average size of a type-θ firm, which is given by $L_0(\theta) = \frac{d\Lambda_0(\theta)/d\theta}{\gamma(\theta)}$, is a continuous function of θ.

3 The constrained efficient allocation

A social planner constrained by the same search frictions as private agents only has to decide which transition opportunities to take up and which ones to ignore. Recall that opportunities to move from unemployment to employment or from job to job only arise infrequently due to search frictions, while the option to move workers into unemployment is always available. In this paper we will only consider the simple case where the planner never finds it optimal to exercise the latter option, because the value of leisure b^{ω} is sufficiently lower than the productivity of any existing firm in all states and/or because employed search is sufficiently effective relative to unemployed search. The important question of employer entry and exit is left for future research.

The constrained efficient allocation is then simple enough to characterize. Let $L^{\star}(\theta)$ denote the density of workers efficiently allocated to a typical type-θ firm and $\Lambda^{\star}(\theta) = \int_{\underline{\theta}}^{\theta} L^{\star}(x)\, d\Gamma(x)$ the corresponding cumulated measure up to θ, so that the unemployment rate is $u^{\star} = 1 - \Lambda^{\star}\left(\overline{\theta}\right)$. Under our assumption that productivity is always large enough that employment in any firm is always socially superior to unemployment, the planner will take up any opportunity to move an unemployed worker into employment, and unemployment evolves according to $u^{\star\prime} = \delta^{\omega'}(1 - u^{\star}) + \left(1 - \lambda_0^{\omega'}\right) u^{\star}$, where primes denote next-period values. Moreover, the planner always seeks to move employed workers from less productive toward more productive firms. Appealing to a Large Numbers approximation, this induces the following simple Markov process for the evolution of efficient firm size:

$$L^{\star}(\theta)' = L^{\star}(\theta)\left(1 - \delta^{\omega'}\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right) + \lambda_0^{\omega'}\left(1 - \Lambda^{\star}\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right)\Lambda^{\star}(\theta). \quad (1)$$

Given new aggregate state ω', of the $L^\star(\theta)$ workers initially employed by this firm, a fraction $\left(1-\delta^{\omega'}\right)$ are not separated exogenously into unemployment. Of these survivors, a fraction $\lambda_1^{\omega'}$ receive an opportunity to move to another firm. The planner exercises that option if and only if the new firm is more productive than θ, which is the case with probability $\overline{\Gamma}(\theta) := 1-\Gamma(\theta)$. Initially unemployed workers $u^\star = 1-\Lambda^\star\left(\overline{\theta}\right)$ find jobs with chance $\lambda_0^{\omega'}$. Workers employed at other firms who have not lost their jobs draw with chance $\lambda_1^{\omega'}$ an opportunity to move to the type-θ firm, that the planner exploits if and only if the firm they currently work at has productivity $x<\theta$. The measure of such workers is $\left(1-\delta^{\omega'}\right)\Lambda^\star(\theta)$.

Multiplying through by $\gamma(\theta)$ in (1) and integrating with respect to θ yields:

$$\Lambda^\star(\theta)' = \lambda_0^{\omega'} u^\star \Gamma(\theta) + \left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\Lambda^\star(\theta).$$

To solve this equation forward in time we introduce a time index t. For any initial condition $\Lambda_0^\star(\theta)$ at some (renormalized) initial date 0 such that the aggregate state last switched to ω at time 0 and then remained at ω between 0 and t, or just the given initial $\Lambda_0(\theta)$ when time starts running, the law of motion of $\Lambda^\star$ is a first-order difference equation which solves as:

$$\Lambda_t^\star(\theta) = \left[(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\right]^t \Lambda_0^\star(\theta) + \lambda_0\Gamma(\theta)\sum_{s=1}^{t}\left[(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\right]^{s-1} u_{t-s}^\star. \quad (2)$$

By inspection, $\Lambda_t^\star(\theta)$ is differentiable in θ at all dates t, and one obtains a closed-form expression for the efficient workforce of any type-θ firm:

$$\begin{aligned} L_t^\star(\theta) = \frac{d\Lambda_t^\star(\theta)/d\theta}{\gamma(\theta)} &= (1-\delta^\omega)^t\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)^{t-1}\left[\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)L_0^\star(\theta) + t\lambda_1^\omega\Lambda_0^\star(\theta)\right] \\ &+ \lambda_0^\omega\left\{u_{t-1}^\star + \sum_{s=2}^{t}(1-\delta^\omega)^{s-1}\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)^{s-2}\left[1-\lambda_1^\omega+\lambda_1^\omega s\Gamma(\theta)\right]u_{t-s}^\star\right\}, \end{aligned} \quad (3)$$

where $L_0^\star(\theta)$ was the value of this solution under state $\widehat{\omega}$ at the time of the last state switch from $\widehat{\omega}$ to ω, and again $u_s^\star = 1-\Lambda_s^\star\left(\overline{\theta}\right)$.

If the aggregate state forever stays at ω, the solutions to (2) and (3) converge to:

$$\begin{aligned} \Lambda_\infty^\star(\theta) &= \frac{\delta^\omega\lambda_0^\omega}{\delta^\omega+\lambda_0^\omega}\cdot\frac{\Gamma(\theta)}{1-(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)} \\ &\text{and} \\ L_\infty^\star(\theta) &= \frac{\delta^\omega\lambda_0^\omega}{\delta^\omega+\lambda_0^\omega}\cdot\frac{1-(1-\delta^\omega)(1-\lambda_1^\omega)}{\left[1-(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\right]^2} \end{aligned} \quad (4)$$

which are the familiar steady-state expressions found in the BM model.

As is well known and immediately verifiable from (4), more productive firms are larger and the normalized distribution of employment across firm types $\Lambda_\infty^\star(\theta)/\Lambda_\infty^\star\left(\overline{\theta}\right)$ is increasing, in the sense of stochastic dominance, in λ_1^ω, and decreasing in δ^ω. Intuitively, workers upgrade to higher-θ firms whenever possible, and are able to do so in larger numbers if they receive more opportunities to do so (higher λ_1^ω) or if they get thrown off the job ladder into unemployment less often (lower δ^ω), in which case the stationary employment distribution is even more skewed towards large firms. This comparative statics property is reflected in the dynamic behavior of the employment distribution by firm size, as we illustrate with numerical examples in Section 6. There we assume, as is consistent with empirical evidence on job-to-job quits and job separations, that λ_1^ω is increasing, and δ^ω decreasing in the state of aggregate productivity ω, and, as is suggested by empirical evidence on the size-productivity relationship (and as is in our model's steady state), that more productive firms initially employ more workers. Then, hitting the economy with a randomly drawn sequence of aggregate shocks, we find that large employers are more cyclically sensitive, because they gain workers faster over an aggregate expansion as job upgrading accelerates, and vice versa in a slump. This property of the efficient allocation exactly replicates the new empirical evidence that we document in MPV09.

4 Equilibrium

4.1 Definition

Each firm chooses and commits to an employment contract, namely a state-contingent wage depending on some state variable ζ, to maximize the present discounted value of profits, given other firms' contract offers. The firm is further subjected to an *equal treatment constraint*, whereby it must pay the same wage to all its workers. This is the sense in which we generalize the BM restrictions placed on the set of feasible wage contracts to a non-steady-state environment.[4] Under commitment, such a wage function implies a value V for any worker to work for that firm, which is also a function of the state ζ. For reasons that will become clear shortly, we assume that a contract offered by a firm to its workers is observable only by the parties involved.

Let Z be the (Borel-)measurable set of all histories of play in the game, and $\mathscr{V}_Z$ the set of measurable functions $\left[\underline{\theta}, \overline{\theta}\right] \times Z \to \mathbb{R}$. A behavioral strategy of the contract-posting game

[4]We thus rule out, among other things, wage-tenure contracts (Stevens, 2004; Burdett and Coles, 2003), offer-matching or individual bargaining (Postel-Vinay and Robin, 2002; Dey and Flinn, 2005; Cahuc, Postel-Vinay and Robin, 2006), contracts conditioned on employment status (Carrillo-Tudela, 2009). Note, however, that the model can be generalized to allow for time-varying individual heterogeneity under the assumption that firms offer the type of piece-rate contracts described in Barlevy (2008). In that sense experience and/or tenure effects can be introduced into the model. Shimer (2008) proposes an alternative formulation, which maintains BM's restriction of a constant posted wage, even out of steady state, and delivers a few of the same results in transitional (non-stochastic) aggregate dynamics.

is a function $V \in \mathscr{V}_Z$ such that, when the state of the game is $\zeta \in Z$, each firm $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ offers value $V(\theta, \zeta)$ to all of its workers.

As V is measurable, the c.d.f.

$$F(W \mid \zeta, V) := \int_{\underline{\theta}}^{\overline{\theta}} \mathbb{I}\{V(\theta, \zeta) \leq W\} \, d\Gamma(\theta) \tag{5}$$

is well-defined for every $\zeta \in Z$, $W \in \mathbb{R}$ and $\mathbb{I}$ an indicator function. This is the probability that a randomly drawn firm offers value no greater than W, given history ζ and given that all firms follow strategy V. Let $\overline{F} = 1 - F$ denote the survival function.

Again, let $\Lambda(\theta)$ be the measure of workers currently employed at all firms of productivity up to θ, so $\Lambda\left(\overline{\theta}\right)$ is total employment. For any increasing $\Lambda : \left[\underline{\theta}, \overline{\theta}\right] \to [0, 1]$, $\zeta \in Z$, $W \in \mathbb{R}$, the following c.d.f.

$$G(W \mid \zeta, \Lambda, V) := \frac{1}{\Lambda\left(\overline{\theta}\right)} \cdot \int_{\underline{\theta}}^{\overline{\theta}} \mathbb{I}\{V(\theta, \zeta) \leq W\} \, d\Lambda(\theta) \tag{6}$$

is also well-defined. This is the probability that a randomly drawn worker is currently earning value no greater than W after history ζ.

Given a strategy $V \in \mathscr{V}_Z$ followed by all firms and the resulting F, an unemployed worker earns a value solving:

$$U(\zeta \mid V) = b^{\omega} + \beta \mathbf{E}_{\zeta' \mid \zeta} \left[\left(1 - \lambda_0^{\omega'}\right) U(\zeta' \mid V) + \lambda_0^{\omega'} \int \max \langle v, U(\zeta' \mid V) \rangle \, dF(v \mid \zeta', V) \right], \tag{7}$$

because she collects a flow value b^{ω} and, one period later, when the aggregate state becomes ω', she draws with chance $\lambda_0^{\omega'}$ a job offer from the equilibrium distribution of offered values F, which she accepts if the associated value exceeds that of staying unemployed.

A firm of current size L which posts a value W in state ζ has size zero next period if $W < U$, otherwise, invoking a large numbers approximation, new firm size is:

$$\begin{aligned} L' = \mathscr{L}(\zeta, W \mid V) := {} & L \left(1 - \delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \overline{F}(W \mid \zeta, V)\right) \\ & + \lambda_0^{\omega'} \left[1 - \Lambda\left(\overline{\theta}\right)\right] \mathbb{I}\{W \geq U(\zeta \mid V)\} + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G(W \mid \zeta, V). \end{aligned} \tag{8}$$

After the new aggregate state ω' is drawn, of the measure L of workers currently employed by this firm, a fraction $\left(1 - \delta^{\omega'}\right)$ are not separated exogenously into unemployment. Of these survivors, a fraction $\lambda_1^{\omega'} \overline{F}(W \mid \zeta, V)$ quit because they draw from F an outside offer which gives them a value larger than W. The currently unemployed $1 - \Lambda\left(\overline{\theta}\right)$ find jobs with chance $\lambda_0^{\omega'}$, and accept an offer W from a firm if this is better than unemployment. By random matching, each firm offering more than U receives the same inflow from unemployment. The employed who have not lost their jobs $\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)$ receive an offer with chance $\lambda_1^{\omega'}$, and accept it if the value W they draw is larger than what they were earning before (probability $G(W \mid \zeta, V)$), in which case they quit to this firm offering W.

Adding up, the cumulated firm size evolves as the sum of individual firm sizes on the equilibrium path. For any $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$:

$$\Lambda(\theta \mid \zeta, V)' = \int_{\underline{\theta}}^{\theta} \mathscr{L}(\zeta, V(x, \zeta) \mid V)\, d\Gamma(x) \;\Rightarrow\; \Lambda(\cdot \mid \zeta, V)' := \mathscr{T}(\zeta \mid V). \tag{9}$$

The map $\mathscr{T}$ denotes next period's employment distribution given current state ζ and offer strategy V. The support of Λ is contained in that of Γ, because no worker can be at a firm of type θ if there exists no such firm. By induction, starting from the initial distribution of employment and for every history of the game, Λ has a (possibly nil) Radon-Nikodym derivative $d\Lambda(\theta \mid \zeta, V)/d\Gamma(\theta)$ everywhere in θ, and (9) requires this derivative to be $\mathscr{L}(\zeta, V(\theta, \zeta) \mid V)$. Therefore, F and G also exist at all nodes of the game when firms play the strategy V.

A value strategy $W \in \mathscr{V}_Z$ can also be implemented by a wage strategy $w \in \mathscr{V}_Z$ such that the worker's Bellman equation is solved by W given that all other firms play V: the worker receives the wage and, next period, the expected value of being either displaced, or retained at the same firm, or poached by a higher-paying firm.

$$\begin{aligned} W(\theta, \zeta) = w(\theta, \zeta) + \beta \mathbf{E}_{\zeta' \mid \zeta} \Bigg[& \delta^{\omega'} U(\zeta' \mid V) \\ & + \left(1-\delta^{\omega'}\right) \left(W(\theta, \zeta') + \lambda_1^{\omega'} \int_{W(\theta, \zeta')}^{+\infty} \left[v - W(\theta, \zeta')\right] dF(v \mid \zeta', V) \right) \Bigg] \end{aligned} \tag{10}$$

We are now going to define an equilibrium of the contract-posting game. Each firm plays a game against other firms as well as vis-à-vis its current and prospective workers. Workers act sequentially, as they are always free to quit. Firms follow a behavioral strategy V (a value policy) that must be a best-response against other firms at any node ζ of the game, including those reached with probability zero on the equilibrium path. For example, a firm may find itself losing more workers than predicted by current equilibrium play. This requires specifying a consistent belief assessment. The constraint of delivering the value to the workers once hired is binding because, after hiring a worker with a promise of W, the firm would like to renege and to squeeze the worker against the participation constraint $W = U$. The reputational underpinnings of the firm's commitment power have been explored in the wage-posting literature (Coles, 2001). As is standard, behavioral strategies in the extensive form dynamic game generate strategy profiles of the equivalent static, strategic form game where each firm chooses a map once and for all at time 0.

Our first task is to find the state space Z on which equilibrium strategies can be conditioned. By assumption, past play by other firms is unobservable, hence cannot be part of Z. For the same reason, and because it is small, a firm takes its competitors' behavior (the distributions F and G) as given when choosing a strategy: its own deviations cannot be detected and be subject to retaliation, so its actions cannot affect the distribution of offers in the economy. Each firm can only observe the set Q of public histories of $\{\omega, F, G, \Lambda\}$.

We look for the smallest subset $Z \subseteq Q$ which is sufficient for Q. For every $\zeta \in Z$ and $V(\cdot, \zeta)$, the current offer distribution $F(\cdot \mid \zeta, V)$ is uniquely determined from (5), so it contains no independent information about ζ. The same is true, from (6), of $G(\cdot \mid \zeta, V)$, given ζ, V and Λ. Next, each individual firm takes the strategy V chosen by others as given, whether or not this firm is maximizing, given ζ. Therefore, for every V, a firm can calculate the history of Λ based only on the history of ω. That is, each firm takes the path of employment at other firms Λ as an exogenous stochastic process. Hence, for every value-offering strategy defined on the history of ω and Λ, there exists an equivalent value-offering strategy defined on the history of ω only, given Λ_0, which produces the same payoff relevant variables for firm θ. For the purpose of calculating firm θ's best response, the history of ω is sufficient for the history of Λ.

The only other independent piece of information that is relevant to a firm's profit maximization is own size L, that is directly controlled by the firm and has a direct impact on the firm's continuation payoffs. Because the history of own size $\{L_s\}_{s=1}^{t}$ is private information, it cannot affect values offered by other firms. Hence only current size L_t can affect the firm's best response, because of its direct impact on profits. We conclude that the only strategically relevant history for a firm can be $\zeta_t = \{\omega_1, \cdots, \omega_t, L_t\}$. Clearly, past values of ω cannot be ruled out of the state space Z, as they are exogenous and public events that firms can use to coordinate actions, for example as a public randomization device, although they are no longer payoff-relevant given the Markov evolution of ω.

Definition 1 *A* SEQUENTIAL EQUILIBRIUM *of the contract posting game is a measurable function $V \in \mathscr{V}_Z$ of, and a set of consistent beliefs over the set Z of histories of the aggregate productivity state and current size, such that V maximizes the present discounted value of profits, given that all other firms also play V and given beliefs. Formally, at least one solution $w \in \mathscr{V}_Z$ to* (10) *with $W = V$ also solves:*

$$w(\theta, \zeta) = \arg\max_{\widetilde{w}} \mathbf{E}\left[\sum_{t=0}^{+\infty} \beta^t \left(\omega_t \theta - \widetilde{w}(\theta, \zeta_t)\right) \mathscr{L}(\zeta_{t-1}, W \mid V) \mid \zeta_0 = \zeta\right],$$

where $\Lambda_t(\theta) = \int_{\underline{\theta}}^{\theta} \mathscr{L}(\zeta_{t-1}, V(x, \zeta_{t-1}) \mid V)\, d\Gamma(x)$, and F, G, U, W are defined by (5), (6), (7), (10) *with $\zeta = \zeta_t$.*

The equilibrium strategy V is a fixed point in the usual game-theoretic sense: if all firms follow V and workers act optimally, then given the implied evolution of the cross-section distributions of values offered F and earned G and of the value of unemployment U, each firm θ's best response is to follow the same strategy $W = V$. The strategy is specified for every possible public history, and individual deviations are unobservable. Thus, a set of consistent beliefs is that all other firms will play according to V after any observed history, on and off the equilibrium path.

4.2 Markov perfect equilibrium

Because the history of aggregate productivity is too large a space to be tractable, we look for equilibria in strategies that depend only on current values of payoff-relevant variables. From our discussion, it is clear that

$$\hat{\zeta} = \{\theta, L, \omega', \Lambda\} \tag{11}$$

is both the smallest and largest such state vector on which equilibrium strategies can depend. If all firms condition their current offers on these four objects in $\hat{\zeta}$, then from Definition 1 of equilibrium so should each firm in its best response. Let $\mathscr{V}_{\hat{Z}}$ be the space of measurable functions $\hat{Z} \to \mathbb{R}$. Then we focus on:

Definition 2 *A* Markov perfect equilibrium *of the contract posting game is a sequential equilibrium V in the set $\mathscr{V}_{\hat{Z}} \subset \mathscr{V}_{Z}$, a measurable function of $\hat{\zeta}$ defined in* (11).

Making strategies independent of past values of aggregate productivity comes at the cost of introducing in the state the current distribution of employment Λ. This is also an infinitely dimensional object, but it turns out to be much more tractable, as we will see next.

From now on, we let the new value distributions $F(\cdot \mid \omega', \Lambda)$ and $G(\cdot \mid \omega', \Lambda)$, firm size $\mathscr{L}(L, \omega', \Lambda, W)$, employment distribution $\mathscr{T}(\omega', \Lambda)$, and value of unemployment $U(\omega', \Lambda)$ be defined as in (5) - (9), with the Markov state in (11) replacing ζ. Notice that only new firm size $\mathscr{L}$ depends on L; the other objects only depend on the aggregate components of the state, ω' and Λ, that each firm takes as given stochastic processes on and off the equilibrium path. That is, $\hat{\zeta}$ contains only one endogenous (to the firm) state variable, its own size L.

5 Equilibrium characterization

5.1 The firm's contract-posting problem: recursive formulation

We look for a Markov perfect equilibrium of the contract-posting game. Suppose all other firms offer a value $V(\theta, L, \omega', \Lambda)$ which depends on own productivity θ, beginning-of-period own size L and distribution of employment Λ, and new state of aggregate productivity ω'. Then, by inspection of the firm's sequential profit maximization problem, these four objects are sufficient to pin down the firm's best response and evolve according to a Markov process. Therefore, it is natural to seek a recursive formulation of the firm's problem. As is standard in the contracting literature (Spear and Srivastava, 1987), the firm's sequential contracting problem is equivalent to a recursive problem, in which the firm takes the value currently promised to its workers as a state variable, and faces a promise-keeping constraint.

We fix the strategy of other firms V and omit it from the notation for simplicity. The firm can always guarantee itself zero flow profits by making the participation constraint $W(\omega') \geq U$ bind and dismissing all workers, so offering any value lower than U is equivalent to an offer $W(\omega') = U$. Using the law of motions of own employment and of the aggregate employment distribution, the firm solves:

$$\Pi\left(\theta, L, \omega, \Lambda, \overline{V}\right) = \sup_{w, W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))} \Big\langle (\omega\theta - w) L \\ + \beta \int_{\Omega} \Pi\left[\theta, \mathscr{L}\left(L, \omega', \Lambda, W(\omega')\right), \omega', \mathscr{T}(\omega', \Lambda), W(\omega')\right] H(d\omega' \mid \omega) \Big\rangle \quad (12)$$

subject to a Promise-Keeping (PK) constraint to deliver the promised $\overline{V}$:

$$\overline{V} = w + \beta \cdot \int_{\Omega} \Big\{ \delta^{\omega'} U\left(\omega', \mathscr{T}(\omega', \Lambda)\right) + \left(1 - \delta^{\omega'}\right) \\ \cdot \left[\left(1 - \lambda_1^{\omega'} \overline{F}\left(W(\omega') \mid \omega', \Lambda\right)\right) W(\omega') + \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v \mid \omega', \Lambda)\right] \Big\} H(d\omega' \mid \omega), \quad (13)$$

where the continuation value on the RHS comes from (10) after a small algebraic manipulation. In (12), given the timing of events, the firm collects flow revenues, equal to per worker productivity $\omega\theta$ times firm size L, then pays the flow wage w to each worker, then observes the new state of aggregate productivity ω', and finally chooses the continuation contract (promised value) $W(\omega')$, so that wage and continuation values deliver at least the current expected value $\overline{V}$ to the workers.

Notice that at time 0 the firm could extract full rents by offering $w = -\infty$, because it is "too late" for the initial workers to quit. To avoid this pathological outcome, we let the initial wage be chosen according to some bargaining procedure that splits rents from the contract and leaves the firm a non-negative cut, so the recursive formulation (12)-(13) only applies from period $t = 1$ on. Therefore, $\overline{V} \geq U(\omega, \Lambda)$ is always guaranteed, because $\overline{V}$ is the value promised a period before under the worker participation constraint $W(\omega') \geq U$.

5.2 Properties of F and G in equilibrium

We begin by establishing that the distributions of offered and accepted worker values, F and G, must satisfy certain general properties in equilibrium, which parallel similar properties of the corresponding wage distributions in the original BM model.

Proposition 1 (F and G are atomless) *In equilibrium F and G must be atomless at all dates and in all states, with their common support being compact and convex.*

To see why there cannot be an atom in F or G, observe that, by the equal treatment constraint, if F had an atom at some value W, then so would G. But an atom in G would open the way to a profitable deviation, as in BM. A firm that is part of the atom that offers the same W in some state could deviate, offer an epsilon more, win the competition for employed workers against all other competitors offering W, and poach an additional positive measure of workers at a negligible marginal cost. This deviation is unprofitable only if the firm was already offering its workers so much as to break even in expected present discounted terms. But then a deviation toward offering, e.g., $W = U$ in all states is profitable as all unemployed workers accept this offer and stay for a while, generating strictly positive profits for all but the zero measure of firms with marginal type that break even with $W = U$.

To see why the support of offered and paid values is convex, observe that if there was a gap then the lower and upper bounds of this gap would generate the same hiring and retention, so the same firm size, but the upper bound would cost the firm more in terms of wages, so no firm would post such an upper bound. To see why the support is compact, observe that $\overline{W} = \max \omega\theta/(1-\beta)$ is a natural upper bound to the offered value: the firm can always do weakly better by offering less than $\overline{W}$, as it can hope to make some profits. So the support is a convex and bounded subset of $\mathbb{R}_+$, which we can therefore take to be compact WLOG.

The properties stated in Proposition 1 will simplify our further characterization of equilibrium, to which we now turn.

5.3 An equivalent unconstrained recursive formulation

We define the joint value of the firm-worker collective as:

$$S = \Pi + \overline{V}L.$$

Next solving for the wage from (13) and replacing it into the firm's Bellman equation (12) we see that the joint value function S solves:

$$S(\theta, L, \omega, \Lambda) = \omega\theta L + \beta \int_{\Omega} \Bigg\{ \delta^{\omega'} U(\omega', \mathscr{T}(\omega', \Lambda)) L$$

$$+ \max_{W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))} \Bigg\langle S(\theta, \mathscr{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathscr{T}(\omega', \Lambda)) + L\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v \mid \omega', \Lambda)$$

$$- W(\omega') \left(\lambda_0^{\omega'} \left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G\left(W(\omega') \mid \omega', \Lambda\right)\right) \Bigg\rangle \Bigg\} H(d\omega' \mid \omega). \quad (14)$$

Notice that the promised value $\overline{V}$ does not appear as an argument of S in the above equation: inspection reveals that the DP problem in (14) is independent of $\overline{V}$. Along the optimal path, the level of current promised utility $\overline{V}$ only affects the distribution of payoffs between the firm and its workers, not their overall level S. To see why, observe that the workers' turnover decisions only depend on continuation values $W(\omega')$ promised by the firm, therefore the same applies to firm continuation profits $\Pi(\omega')$. The firm thus chooses $W(\omega')$ to maximize $\Pi(\omega')$ independently of the currently promised value $\overline{V}$. Then, to deliver $\overline{V}$ as promised without distorting the optimally set future turnover, the firm adjusts the current wage w.

The optimal policy solving the unconstrained DP problem (14) also solves (12) subject to (13). We therefore focus on the analysis of the simpler problem (14). An equilibrium is a solution V that coincides with the one followed by the other firms. To find the equilibrium, we proceed as follows. First, we show that under certain sufficient conditions a best response value to any strategy followed by all other firms must be strictly increasing in own productivity θ and size L. Next, in that smaller set of monotonic functions we construct an equilibrium.

5.4 Rank-Preserving Equilibrium (RPE)

While solving for equilibrium directly is an intractable problem because the size distribution of firms Λ is an infinitely-dimensional state variable, we can still define a tractable and natural class of equilibria, which have the following property. Let $L(\theta)$ denote employment size of a type-θ firm along the equilibrium path, i.e. the size attained by that firm given the initial size distribution at date 0 and given that all firms have played the equilibrium strategy from date 0 up to the current date. Then:

Definition 3 *An equilibrium is* Rank-Preserving *(RP) if a more productive firm always pays its workers more: $\theta \mapsto V(\theta, L(\theta), \omega, \Lambda)$ is increasing in θ.*

A direct consequence of the above definition is that in a Rank preserving Equilibrium (RPE) workers rank their preferences to work for different firms according to firm productivity at all dates. The following two properties thus hold true at all dates under the RP assumption: the proportion of firms that offer less than $V(\theta, L(\theta), \omega, \Lambda)$ is simply that proportion of firms that are less productive than θ

$$F\left(V\left(\theta, L\left(\theta\right), \omega', \Lambda\right) \mid \omega', \Lambda\right) \equiv \Gamma\left(\theta\right), \tag{15}$$

and the number of employed workers who earn a value that is lower than that offered by θ equals employment at firms less productive than θ:

$$\Lambda\left(\overline{\theta}\right) G\left(V\left(\theta, L\left(\theta\right), \omega', \Lambda\right) \mid \omega', \Lambda\right) = \Lambda\left(\theta\right). \tag{16}$$

As we will see those restrictions will decisively simplify the calculations involved in solving for equilibrium in the stochastic model. Moreover, the RP property is theoretically appealing for at least two more reasons. First, it parallels a well-known property of the static equilibrium characterized by BM, which is to have a unique equilibrium where workers rank firms according to productivity. Second, RPE feature constrained-efficient labor reallocation at all dates: if workers consistently rank more productive firms higher than less productive ones, then job-to-job moves will always be up the productivity ladder. That is, in any RPE $L(\theta) = L^{\star}(\theta)$ and the allocation is unique.

It is therefore natural to ask how general Rank-Preserving Equilibria are. We now show that under some weak sufficient conditions on the initial size distribution of employment, all Markov equilibria must have this property. This is the central result of the paper. We assume that Ω is finite only for simplicity of exposition and proof, to avoid dealing with measurability issues, but nothing conceptually depends on this restriction.

Proposition 2 (Ranked initial firm size implies rank-preserving equilibrium) *Assume Ω is finite. If at the initial date 0 the initial state of the economy is such that L_0 is non-decreasing in θ (i.e. higher-θ firms start out no smaller), then any symmetric Markov Perfect equilibrium of the dynamic value-posting game is necessarily Rank-Preserving, and the initial ranking of firms' relative sizes is maintained on the equilibrium path. If Γ is degenerate and firms are equally productive, then the same conclusion holds and initially larger firms offer more and remain larger on any equilibrium path.*

Although the proof, in Appendix A, is technically quite involved, the proposition has a simple economic intuition. In BM's steady-state model, more productive firms offer higher wages due to a single-crossing property of their steady state profits, which in turn reflects two very basic economic forces. First, a higher wage implies a larger firm size, as a more generous offer makes it easier to poach workers and to fend off competition. Second, a larger firm size is more valuable to a more productive firm, because each worker produces more. Therefore, by a simple monotone comparative statics argument, it must be the case that more productive firms offer more, employ more workers, and earn higher profits. Simply put, a productive firm can afford paying more, and is willing to do so to attract workers, because its opportunity cost of not producing is higher. Key to this argument is the fact that firm size is an endogenous object, and BM look for an appropriate firm size distribution which guarantees a stationary allocation.

In our dynamic model, firm size is a state variable, and its *initial* value is a parameter of the model, arbitrarily fixed, not an endogenous object. Therefore, in order to get a start on monotone comparative statics, it is sufficient (but not necessary) that the initial size distribution shares the key property of BM's steady state distribution; namely, it is increasing in productivity. In the proof, we essentially invoke a single-crossing property of the maximand (the term in $\langle\cdot\rangle$) in the Bellman equation of the modified but equivalent value-posting problem (14).[5] A more productive firm still wants and can afford to pay more, now in terms of values accruing to workers. If initially (or once) larger, this firm has a further motive to offer more, namely more workers to retain, independently of its productivity. In contrast, the effect of a higher offer on successful poaching from other firms is independent of current size, because of CRS in production. Therefore, the initial ranking of sizes by productivity is preserved throughout, and values offered to workers remain ranked by firm productivity at all points in the future. This condition is only sufficient. We conjecture that it is not necessary. It aligns two separate motives to pay workers more, firm productivity and size, so clearly there is some slack. If firms are equally productive and only differ in their initial size, then only the size motive operates and all equilibria are RP, with no additional conditions.

We stress that this is a characterization result, which neither establishes nor requires existence, let alone uniqueness, of a RPE. Our main result says that, if a Markov Perfect Equilibrium V exists, then V can only be a best response to itself if it is increasing in θ, including the effect of endogenous size on the posted value. So ours is a general monotonicity result, which does not require to either propose or calculate a particular value-offer strategy. In the next section, we show by construction existence and uniqueness of a RPE, which must then be the unique Markov Perfect Equilibrium of the contract-posting game.

To characterize a RPE we need to describe how allocation and prices depend on exogenous states. The allocation is easy because constrained efficient. We already know from Section

[5] In a way similar to that in which Caputo (2003) appeals to single-crossing properties of the Hamiltonian in his analysis of comparative dynamics for *deterministic* optimal control problems.

3 how the size of each firm evolves in equilibrium. Indeed, the same logic applies to any job ladder model in which a similar concept of RPE can be defined. Nothing in the dynamics of $L^\star$ or $\Lambda^\star$ depends on the particulars of the wage setting mechanism, so long as this is such that employed jobseekers move from lower-ranking into higher-ranking jobs in the sense of a time-invariant ranking. Therefore, this model's predictions about everything relating to firm sizes are in fact much more general than the wage- (or value-) posting assumption retained in the BM model. We now turn to supporting prices.

5.5 Properties of optimal contracts in RPE

Equation (3) combined with the assumption that initial firm size, $L_0(\theta)$, is a continuous function of θ (see Section 2) ensures that $L^\star(\theta)$ is a continuous function of θ at all dates in a RPE. With that in mind, we can establish the following additional properties of the joint value function S and worker value function V in a RPE:

Proposition 3 (Differentiability of value functions in RPE) *The following properties hold in a RPE:*

1. *$L \mapsto S(\theta, L, \omega, \Lambda^\star)$ is convex in L and differentiable in L at $L^\star(\theta)$, i.e. $S_L(\theta, L^\star(\theta), \omega, \Lambda^\star)$ exists for all θ. Moreover, $S_L(\theta, L, \omega, \Lambda^\star)$ is continuous in L at $L^\star(\theta)$;*
2. *$\theta \mapsto S_L(\theta, L^\star(\theta), \omega, \Lambda^\star)$ is continuously differentiable in θ;*
3. *$\theta \mapsto V(\theta, L^\star(\theta), \omega, \Lambda^\star)$ is continuously differentiable in θ.*

The proof is in Appendix B. While most of that proof is essentially technical, it begins by establishing continuity of $\theta \mapsto V(\theta, L^\star(\theta), \omega, \Lambda^\star)$, which is intuitive by a simple improvement argument. If V jumps up at some value of θ, the right and left limits of this value at θ generate the same transitions and firm size, but the right limit costs the firm more, and revenues are continuous in θ.

The third statement in Proposition 3 allows us to differentiate (15) and (16) w.r.t. θ:

$$f(V \mid \omega, \Lambda^\star) \cdot \frac{dV}{d\theta} = \gamma(\theta) \qquad \text{and} \qquad g(V \mid \omega, \Lambda^\star) \cdot \frac{dV}{d\theta} = L^\star(\theta)\gamma(\theta). \tag{17}$$

at $V = V(\theta, L^\star(\theta), \omega, \Lambda^\star)$.

These differentiability properties allow the use in (14) of first-order conditions, which, for each state ω', write down as (using the definition of $\mathscr{L}(\cdot)$ again and using subscripts to denote partial derivatives):

$$\lambda_0^{\omega'}\left(1-\Lambda^\star\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^\star\left(\overline{\theta}\right)G\left(W\left(\omega'\right)\mid\omega',\Lambda^\star\right)$$
$$=\left[S_L\left(\theta,\mathscr{L}\left(L^\star\left(\theta\right),\omega',\Lambda^\star,W\left(\omega'\right)\right),\omega',\mathscr{T}\left(\omega',\Lambda^\star\right)\right)-W\left(\omega'\right)\right]$$
$$\times\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\left[L^\star\left(\theta\right)f\left(W\left(\omega'\right)\mid\omega',\Lambda^\star\right)+\Lambda^\star\left(\overline{\theta}\right)g\left(W\left(\omega'\right)\mid\omega',\Lambda^\star\right)\right]-m^{\omega'}\quad(18)$$

where $m^{\omega'}$ is the Lagrange multiplier for the workers' participation constraint $W\left(\omega'\right)\geq U\left(\omega',\mathscr{T}\left(\omega',\Lambda^\star\right)\right)$, and where complementary slackness $m^{\omega'}\left[W\left(\omega'\right)-U\left(\omega',\mathscr{T}\left(\omega',\Lambda^\star\right)\right)\right]=0$ applies. In a RPE, (18) is solved by $W=V\left(\theta,L^\star\left(\theta\right),\omega,\Lambda^\star\right)$. Next, in the firm's problem (14), the Envelope condition w.r.t. firm size writes as:

$$S_L\left(\theta,L^\star\left(\theta\right),\omega,\Lambda^\star\right)=\omega\theta+\beta\int_\Omega\left\{\delta^{\omega'}U\left(\omega',\mathscr{T}\left(\omega',\Lambda^\star\right)\right)+\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{W(\omega')}^{+\infty}vdF\left(v\mid\omega',\Lambda^\star\right)\right.$$
$$\left.+S_L\left(\theta,\mathscr{L}\left(L^\star\left(\theta\right),\omega',\Lambda^\star,W\left(\omega'\right)\right),\omega',\mathscr{T}\left(\omega',\Lambda^\star\right)\right)\left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{F}\left(W\left(\omega'\right)\mid\omega',\Lambda^\star\right)\right)\right\}H\left(d\omega'\mid\omega\right).\quad(19)$$

5.6 Existence and uniqueness of (Rank-Preserving) equilibrium

We now introduce a time index t again. With a slight notational abuse, we denote:

$$V_{t+1}\left(\theta\mid\omega\right):=V\left(\theta,L_t^\star\left(\theta\right),\omega,\Lambda_t^\star\right)\quad\text{and}\quad U_t\left(\omega\right):=U\left(\omega,\Lambda_t^\star\right).$$

We further define the costate variable:

$$\mu_{t+1}\left(\theta\mid\omega\right):=S_L\left(\theta,\mathscr{L}\left(L_t^\star\left(\theta\right),\omega,\Lambda_t^\star,V_{t+1}\left(\theta\mid\omega\right)\right),\omega,\mathscr{T}\left(\omega,\Lambda_t^\star\right)\right),$$

which measures the shadow value to the worker-firm collective of the marginal worker, given the aggregate state, along the equilibrium path. Note that the dependence of V and μ on the state variables $L^\star$ and $\Lambda^\star$ is subsumed into the time index in the above notation, which is licit as those two variables evolve deterministically conditional on ω. Combining (18) and the various restrictions (15), (16), and (17) that hold in a RPE, we obtain the RPE version of the FOC (18):

$$\lambda_0^\omega u_t+\lambda_1^\omega\left(1-\delta^\omega\right)\Lambda_t^\star\left(\theta\right)$$
$$=\lambda_1^\omega\left(1-\delta^\omega\right)\left[\mu_{t+1}-V_{t+1}\right]\left[L_t^\star\left(\theta\right)f\left(V_{t+1}\mid\omega,\Lambda_t^\star\right)+\left(1-u_t^\star\right)g\left(V_{t+1}\mid\omega,\Lambda_t^\star\right)\right]-m_t^\omega$$
$$=2\lambda_1^\omega\left(1-\delta^\omega\right)\frac{L_t^\star\left(\theta\right)\gamma\left(\theta\right)}{dV_{t+1}/d\theta}\left(\mu_{t+1}-V_{t+1}\right)-m_t^\omega,\quad(20)$$

and the RPE version of the Euler equation (19):

$$\mu_t(\theta \mid \omega) = \omega\theta + \beta \int_\Omega \left\{ \delta^{\omega'} U_{t+1}(\omega') + \left(1-\delta^{\omega'}\right) \lambda_1^{\omega'} \int_\theta^{+\infty} V_{t+1}(x \mid \omega')\, d\Gamma(x) \right.$$
$$\left. + \mu_{t+1}(\theta \mid \omega') \left(1-\delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \overline{\Gamma}(\theta)\right) \right\} H(d\omega' \mid \omega), \quad (21)$$

Note that now the shadow marginal value μ only depends on the distribution of employment $\Lambda^\star$ through total employment in all firms of productivity up to θ, $\Lambda^\star(\theta)$ and the corresponding density $L^\star(\theta)\gamma(\theta)$. Both are scalars, and the state reduces from $\widehat{\zeta} = (\theta, L, \omega', \Lambda)$, which is infinite-dimensional due to the relevance of the entire firm size distribution Λ, to the four-dimensional vector $\mathbf{z} = (\theta, L, \omega', \Lambda(\theta))$: in order to make its decisions, the firm only needs to know the mass of employment at less productive firms $\Lambda(\theta)$ and not the entire size distribution Λ.

Finally, a Transversality Condition (TVC) requires that the discounted joint marginal value of adding one worker to the firm vanishes in expectation w.r. to the stochastic path of ω

$$\lim_{t\to\infty} \mathbf{E}\left[\beta^t \mu_t(\theta \mid \omega) L_t^\star(\theta) \mid \mathbf{z}_0\right] = 0. \quad (22)$$

We now assume that $\omega\underline{\theta} \geq b^\omega \geq 0$ for all ω, so that $U \geq 0$, because a worker has always the option of staying unemployed to collect $b^\omega \geq 0$. We also assume that $\lambda_0^\omega - \lambda_1^\omega$ is small enough for every ω that the worker participation constraint never binds in equilibrium.[6] A RPE is then a value V increasing in θ, a shadow value of employment μ, and a value of unemployment U positive and smaller than V, obeying the boundary condition $V_t(\underline{\theta} \mid \omega) = U_t(\omega)$ and solving the FOC (20), the Euler equation (21) and the unemployment Bellman equation (7) given the RPE employment dynamics (1), subject to the TVC (22). Let

$$Q_t(\theta \mid \omega) := \left(\lambda_0^\omega u_t + \lambda_1^\omega (1-\delta^\omega) \Lambda_t^\star(\theta)\right)^2.$$

We can verify by direct substitution that the following value function satisfies the FOC (20) and the boundary condition:

$$V_t(\theta \mid \omega) = \frac{Q_t(\underline{\theta} \mid \omega)}{Q_t(\theta \mid \omega)} U_t(\omega) + \int_{\underline{\theta}}^{\theta} \mu_t(x \mid \omega) \frac{\frac{\partial Q_t}{\partial \theta}(x \mid \omega)}{Q_t(\theta \mid \omega)} dx := \mathbf{T}_V[\mu_t, U_t](\theta \mid \omega).$$

Let $\mathscr{E}_\Theta$ be the space of continuous cdf's over $\Theta = \left[\underline{\theta}, \overline{\theta}\right]$, $\mathscr{F}_{\Theta\times\Omega\times\mathscr{E}}$ be the space of positive functions $\Theta \times \Omega \times \mathscr{E}_\Theta \to \mathbb{R}_+^2$ such that the first component is θ-integrable and the second

[6] Notice that when $\lambda_0 = \lambda_1$ the worker has no reason to decline any offer, and with $\omega\underline{\theta} > b^\omega$ even the least productive firm can hire some unemployed workers and obtain positive profits. We abstract from entry and exit of firms to focus on the poaching competition and job ladder.

component is constant as we vary $\theta \in \Theta$, and $\mathbf{T} = \binom{\mathbf{T}_\mu}{\mathbf{T}_U}$ be the linear function on $\mathscr{F}_{\Theta \times \Omega \times \mathscr{E}}$ defined by

$$\mathbf{T}_\mu [\mu_t, U_t] (\theta \mid \omega) = \int_\Omega \Bigg\{ \delta^{\omega'} U_{t+1} (\omega') + \mu_{t+1} (\theta \mid \omega') \left(1 - \delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \overline{\Gamma} (\theta)\right) \\ + \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_\theta^{\overline{\theta}} \mathbf{T}_V [\mu_{t+1}, U_{t+1}] (x \mid \omega') \, d\Gamma (x) \Bigg\} H (d\omega' \mid \omega)$$

$$\mathbf{T}_U [\mu_t, U_t] (\omega) = \int_\Omega \left\{ \left(1 - \lambda_0^{\omega'}\right) U_{t+1} (\omega') + \lambda_0^{\omega'} \int_{\underline{\theta}}^{\overline{\theta}} \mathbf{T}_V [\mu_{t+1}, U_{t+1}] (x \mid \omega') \, d\Gamma (x) \right\} H (d\omega' \mid \omega)$$

Note that by definition of these mappings, $\mathbf{T}$ preserves positivity of its arguments and the second component is independent of θ, so that $\mathbf{T}$ maps $\mathscr{F}_{\Theta \times \Omega \times \mathscr{E}}$ into itself, whenever the function is well defined (the integrals exist).

Then a RPE is a solution $\binom{\mu}{U} \in \mathscr{F}_{\Theta \times \Omega \times \mathscr{E}}$ of

$$\begin{pmatrix} \mu \\ U \end{pmatrix} = \begin{pmatrix} \omega\theta \\ b^\omega \end{pmatrix} + \beta \mathbf{T} \begin{pmatrix} \mu \\ U \end{pmatrix}, \tag{23}$$

which satisfies the TVC (22) and has $0 \leq U \leq \mathbf{T}_V [\mu, U] \leq \mu$ and $\mathbf{T}_V [\mu, U]$ increasing in θ.

We are now in a position to prove the following result:

Proposition 4 (Uniqueness and Existence) *There exists at most one equilibrium, which is Rank-Preserving. If it exists, the optimal contract in this unique RPE is the wage policy that pays the worker a value* $\mathbf{T}_V [\mu^\star, U^\star]$ *where:*

$$\begin{pmatrix} \mu^\star \\ U^\star \end{pmatrix} (\theta \mid \omega) := \lim_{n \to \infty} \sum_{j=0}^{n} \beta^j \mathbf{T}^j \begin{pmatrix} \omega\theta \\ b^\omega \end{pmatrix} (\theta \mid \omega) . \tag{24}$$

Existence is guaranteed under the sufficient conditions $\omega\underline{\theta} \geq b^\omega$ *and* $\lambda_0^\omega \leq \lambda_1^\omega (1 - \delta^\omega) \; \forall \; \omega$.

The proof, in Appendix C, simply proceeds through forward substitution and induction, and establishes also that this limit exists. While we have not been able to derive conditions on parameters that are both necessary and sufficient for equilibrium existence (i.e. for (μ^*, U^*) to be a RPE), this is not an issue in applications. In fact, we proved that there is only one possible equilibrium set of contracts, that we can compute and then check ex post whether in fact it satisfies all equilibrium conditions.

6 Practical implementation of RPE

6.1 A strategy to solve for stochastic RPE

We now show how to numerically "solve" for the RPE, by which we mean simulate the dynamic paths of the distributions of employment and wages across firms, given an initial state and a subsequent realization of a sequence of aggregate shocks. Our main goal in this final section is to demonstrate that the conceptual simplicity of the model's equilibrium translates into computational tractability. This is important because it extends the scope of the search and contract posting approach to aggregate labor market analysis. This class of models have provided valuable insights into the causes of wage dispersion, worker turnover, individual wage dynamics, and the firm size/wage relationship. So far, however, their application to business cycle analysis has been considered an interesting but intractable problem. To fill this gap, we construct a numerical example based directly on the model of previous sections. This example not only illustrates how to compute equilibrium, but also highlights quantitative successes and failures of this simple version of the model. As we discuss in the Conclusions, full quantitative success requires extending the model on a few dimensions, in a way that we do now expect to overturn our main theoretical results. We leave a full quantitative analysis for future research.

For the sake of illustration, we focus on the case where the aggregate state can take on two values, $\Omega = \{\omega, \omega'\}$, with conditional switching probabilities σ^{ω} and $\sigma^{\omega'}$. The Euler equation — or Envelope condition — (21) becomes:

$$\begin{aligned}\mu_t(\theta \mid \omega) = \omega\theta + \beta\sigma^{\omega}\Bigg\{ &\delta^{\omega'} U_{t+1}(\omega') + \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'} \int_{\theta}^{+\infty} V_{t+1}(x \mid \omega')\, d\Gamma(x) \\ &+ \mu_{t+1}(\theta \mid \omega')\left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\Bigg\} \\ + \beta(1-\sigma^{\omega})\Bigg\{ &\delta^{\omega} U_{t+1}(\omega) + (1-\delta^{\omega})\lambda_1^{\omega} \int_{\theta}^{+\infty} V_{t+1}(x \mid \omega)\, d\Gamma(x) \\ &+ \mu_{t+1}(\theta \mid \omega)(1-\delta^{\omega})\left(1-\lambda_1^{\omega}\overline{\Gamma}(\theta)\right)\Bigg\}.\end{aligned}$$

Taking derivatives w.r. to θ on both sides, which is possible by Proposition 3:

$$\begin{aligned}\frac{\partial\mu_t}{\partial\theta}(\theta \mid \omega) = \omega + \beta\sigma^{\omega}\left(1-\delta^{\omega'}\right)&\left\{\lambda_1^{\omega'}\gamma(\theta)\,\pi_{t+1}(\theta \mid \omega') + \frac{\partial\mu_{t+1}}{\partial\theta}(\theta \mid \omega')\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\right\} \\ + \beta(1-\sigma^{\omega})(1-\delta^{\omega})&\left\{\lambda_1^{\omega}\gamma(\theta)\,\pi_{t+1}(\theta \mid \omega) + \frac{\partial\mu_{t+1}}{\partial\theta}(\theta \mid \omega)\left(1-\lambda_1^{\omega}\overline{\Gamma}(\theta)\right)\right\}, \quad (25)\end{aligned}$$

where $\pi_t(\theta \mid \omega) := \mu_t(\theta \mid \omega) - V_t(\theta \mid \omega)$ denotes the shadow value to the *firm* of the marginal worker. Together with the FOC for an interior solution of the promised value:

$$\frac{\partial\pi_{t+1}}{\partial\theta}(\theta \mid \omega) = \frac{\partial\mu_{t+1}}{\partial\theta}(\theta \mid \omega) - \frac{2\lambda_1^{\omega}\Lambda_t^{\star\prime}(\theta)(1-\delta^{\omega})}{\lambda_0^{\omega}u_t + \lambda_1^{\omega}(1-\delta^{\omega})\Lambda_t^{\star}(\theta)}\pi_{t+1}(\theta \mid \omega) \quad (26)$$

this gives a system of four functional equations in $\pi_t(\theta \mid \omega)$, $\partial\mu_t(\theta \mid \omega)/\partial\theta$, all functions of θ and t, a pair for each value of ω.

The main difficulty in solving this system lies in the dependence of $\partial\mu_t(\theta \mid \omega)/\partial\theta$ on $\partial\mu_{t+1}(\theta \mid \omega')/\partial\theta$ and $\pi_{t+1}(\theta \mid \omega')$, that is on the jump in the shadow marginal values of one worker, both to the firm (π) and to the collective (μ), caused by the possible occurrence of an aggregate state switch next period. To get around this problem, we can approximate that "jump term" by a known function J (e.g. polynomials) of the state variable $(\theta, L^\star(\theta), \omega, \Lambda^\star(\theta))$ depending on a finite vector of unknown coefficients, $\mathbf{a}$, which can be determined iteratively by successive approximations. Thus the proposed simulation protocol is akin to a projection method to solve the system of PDEs that characterize equilibrium. Its specific feature is that projection is only used to approximate the jumps in π and $\partial\mu/\partial\theta$ caused by aggregate shocks, the rest of the system being solved "exactly". Practical details of the algorithm are given in Appendix D.

Simulation of the model then goes as follows: first, we pick an initial state of the economy $(\omega_0, \Lambda_0(\cdot), L_0(\cdot))$ and simulate a path of ω. Second, given the simulated path of ω and the initial state of the economy, simulate the associated paths of $\Lambda^\star(\cdot)$ and $L^\star(\cdot)$ as per equations (2) and (3). Third, given the previously simulated objects, we solve (25) and (26) subject to (22) using the algorithm sketched above and described more completely in Appendix D. Completion of those three steps produces a solution for $\{\Lambda^\star, L^\star, \partial\mu_t/\partial\theta, \pi_t\}$ over some initially chosen time interval $t \in \{0, \cdots, T\}$ given any simulated sequence of aggregate states. Wages are finally retrieved from:[7]

$$w_t(\theta \mid \omega) = \omega\theta - \pi_t(\theta \mid \omega) + \beta\sigma^\omega\left(1 - \delta^{\omega'}\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\pi_{t+1}(\theta \mid \omega') \\ + \beta(1 - \sigma^\omega)(1 - \delta^\omega)\left(1 - \lambda_1^\omega\overline{\Gamma}(\theta)\right)\pi_{t+1}(\theta \mid \omega).$$

6.2 Calibration and Simulation Results

We now illustrate the quantitative properties of the model using the following calibration. First, all scalar parameters are given values as indicated in Table 1. Next, the sampling distribution of firm types is calibrated following the Bontemps, Robin and Van den Berg (2000) estimation procedure in such a way that the predicted steady-state distribution in the high productivity state fits the business sector wage distribution observed from the CPS in 2006 (see MPV08 for details).

The aggregate productivity shifter ω is normalized to one in the low aggregate state and assumed to be 1% higher in the high state, so that aggregate labor productivity will exhibit fluctuations of roughly $\pm 3\%$ around its trend (see Figure 3 below), which is in same order

[7]Note that this equation also features a "jump term" — i.e., $w_t(\theta \mid \omega)$ depends on future values of π in both aggregate states. We approximate that jump term in the same fashion as we do the jump term in the differentiated Euler equation (25).

	ω	λ_0^ω	λ_1^ω	δ^ω	σ^ω	β
Expansion	1.01	0.206	$0.25 \cdot \lambda_0^\omega$	0.011	0.0238	0.9959
Contraction	1	0.159	$0.25 \cdot \lambda_0^\omega$	0.012	0.0238	0.9959

Table 1: Model Calibration

of magnitude as Robin's (2009) estimates. The high-state value of λ_0 is the monthly job finding rate from the Current Population Survey, averaged over periods during which the U.S. unemployment rate was below its HP trend since 1967. The low-state value of λ_0 is defined in the same way, the average over periods of above-trend unemployment. The job loss rate δ is also constructed in a similar fashion, using averages of the monthly rate of separations out of employment.[8] The arrival rate of offers to employed job seekers, λ_1, is calibrated as a constant fraction of λ_0 which produces a share of job-to-job transitions in total separations of about 50%, as is typically observed (e.g. Hall, 2005). The aggregate productivity process is calibrated in such a way that unemployment is below (resp. above) trend half of the time, with the duration of an average period of below-trend (resp. above-trend) unemployment equalling 3.5 years (as is the case for the monthly U.S. unemployment rate on average since 1948). The discount factor β corresponds to an annual discount rate of 5 percent.

We also simplify the model slightly for computational purposes by imposing a constant binding minimum wage $\underline{w}$, so that the lowest-θ firm in the market will always offer a constant wage of $\underline{w}$. This binding minimum wage substitutes the more complex boundary condition $V(\underline{\theta} \mid \omega) = U(\omega)$ in the system characterizing equilibrium. Finally, we abstract from entry or exit of firms over the cycle. In particular the lower bound of the support of firm-specific productivity components, $\underline{\theta}$, is normalized at 1 in both aggregate productivity states, and so is the minimum wage, $\underline{w}$. This, together with low enough values of leisure b^ω in all states to keep workers from ever wanting to quit into unemployment, implies that even the least productive firm in the economy remains viable in bad times, so that aggregate downturns do not cause firms to leave the market.

Figures 1-4 illustrate the output of a representative simulation. The economy's initial state is set to the low-ω steady state, and at time 0 is hit by a positive aggregate productivity shock: the economy starts off in an expansionary phase, just out of a very (infinitely) long recession. The simulated series then cover a period of 30 years spanning several expansions and recessions. On all plots the cycle is materialized by the dashed line which represents the unemployment rate (in deviations from its mean and rescaled for legibility).

[8] The worker flow data used for this calibration was constructed by Robert Shimer (`http://sites.google.com/site/robertshimer/research/flows`).

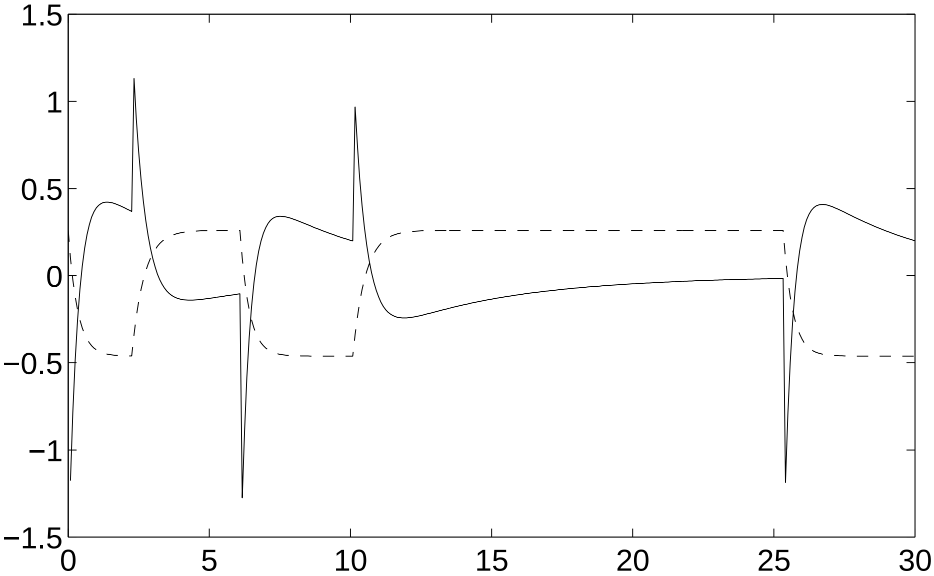

Fig. 1: Firm size growth rate differential (solid) and unemployment rate (dashed)

The results do not differ qualitatively from those of the deterministic transition dynamics analyzed in MPV08. Most remarkably, as documented in MPV08 and MPV09, the growth rate differential of employment at (initially) large minus small firms collapses upon a recession and rises slowly through an expansion, as shown on Figure 1 which plots the simulated differential in average growth rates across firm size classes.[9] The cyclical pattern of relative growth by firm size is naturally interpreted in our model as reflecting the slow upgrading of workers to better jobs through on the job search. The governing mechanism is the following. At the onset of an expansion, the many unemployed workers inherited from the previous recession are available for work at any (low) wage that makes work preferable to unemployment. As those unemployed job applicants are willing to accept any offer, the random search process allocates them into firms following the sampling distribution of firm types. However the magnitude of the job finding rate λ_0 is such that the pool of cheap, unemployed job seekers dries out quickly. To keep expanding, firms begin to raise wages to poach labor from their competitors. But as argued earlier, the BM single-crossing argument applies: more productive firms are both able and willing to raise wages further than their less productive competitors, as their opportunity cost of not producing is higher, especially so

[9]Formally, the average growth rate differential is defined as: $\frac{\Lambda^\star_{t+1}(\overline{\theta})-\Lambda^\star_{t+1}(\theta_\ell)}{\Lambda^\star_t(\overline{\theta})-\Lambda^\star_t(\theta_\ell)} - \frac{\Lambda^\star_{t+1}(\theta_s)}{\Lambda^\star_t(\theta_s),}$ where θ_ℓ and θ_s are thresholds defining the groups of "large" and "small" firms, respectively. Here the group of large firms is defined as the top tercile of the (high-) steady-state distribution of firm sizes among workers, while small firms are the bottom tercile of that distribution. According to data on the distribution of firm sizes from the Business Dynamics Statistics those thresholds roughly correspond to firms of over 1000 and less than 50 employees, respectively, which are the thresholds used in MPV09.

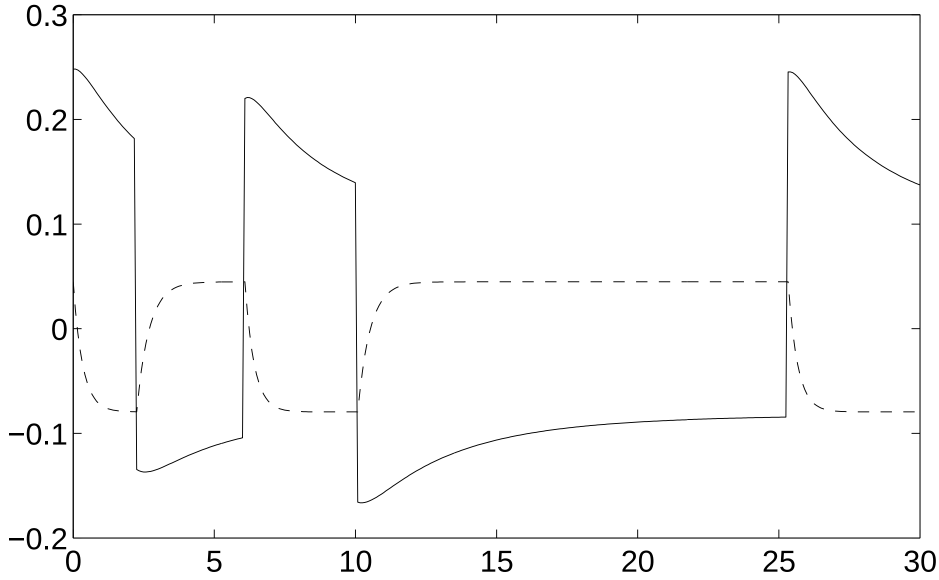

Fig. 2: Job-to-job transition rate (solid) and unemployment rate (dashed)

in an aggregate expansion. Thus workers gradually select themselves into more productive, better paying jobs — at the speed permitted by search frictions, which is determined by λ_1 — and more productive firms grow relatively faster and become relatively larger.[10] As a consequence, as the expansion progresses, the distribution of employment becomes increasingly skewed toward more productive, larger firms. Therefore, when the recession hits, while all firms have too many workers, larger firms have relatively more excess employment. The distribution of firm sizes must then ebb back toward smaller firms, so in net terms large firms shed proportionally more workers to reduce their share. Finally, note that the same intuition explains the negative correlation between the rates of unemployment and job-to-job transition predicted by our model (Figure 2).

The cyclical upgrading of workers also strongly propagates and amplifies the effects of the aggregate labor productivity shock on measured average labor productivity (Figure 3). Labor being a quasi-fixed factor in this model, average output per worker jumps upon impact of an aggregate, proportional productivity shock by $\pm 1\%$, the calibrated magnitude of the aggregate shock (see Table 1). After that initial jump, output per worker continues to adjust in the direction of the shock in a smooth and quasi-monotonic fashion. This smooth adjustment is driven by the slow movements of workers up and down the job ladder: for example in an expansion, labor is slowly reallocated toward more productive firms, which

[10]Indeed an immediate implication of the RP property of equilibrium is that a firm's rank in the distribution of firm sizes is the same as that firm's rank in the productivity distribution or in the distribution of offered worker values. As such our model provides a theoretical justification for the use of firm size as a proxy for firm productivity — as does any job-ladder model whose equilibrium has the RP property.

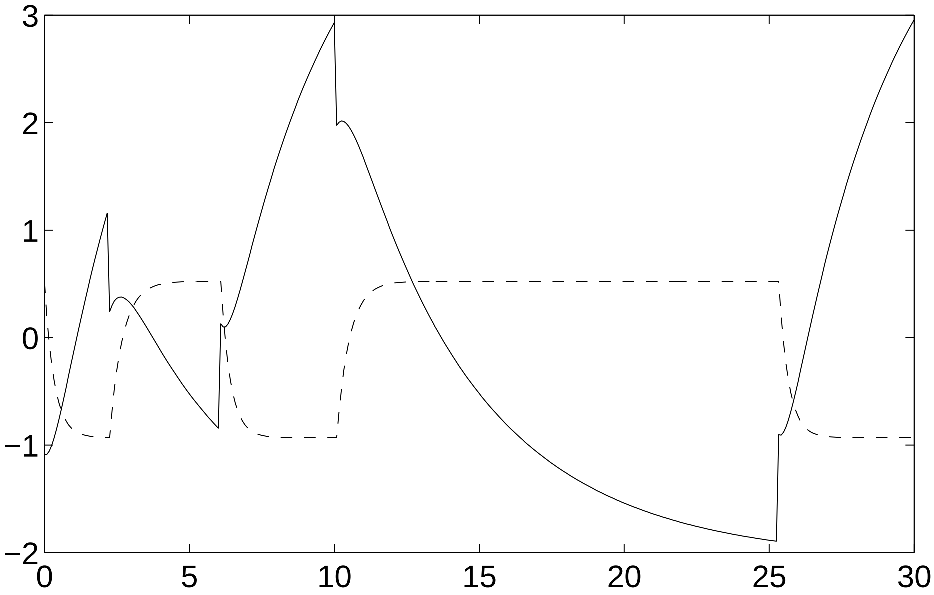

Fig. 3: Mean output per worker (solid) and unemployment rate (dashed)

increases average labor productivity beyond the initial increase in the aggregate productivity shifter through a composition effect. Moreover, because the driving force is the reallocation of workers into better jobs, the speed at which that composition effect drives average productivity up is limited by the speed at which employed workers are able to find better jobs, before they lose the ones they have. So the key parameter that determines the extent of propagation of aggregate shocks in this model is the ratio λ_1/δ, a standard measure of the extent of search frictions in wage-posting models.[11]

The model predicts procyclical wages in the sense that on average steady-state wages are lower in the low aggregate state than in the high aggregate state. We also observe that the wage jumps in a direction opposite to labor productivity when aggregate shocks hit. For example, due to wage backloading, wages drop on impact following a positive aggregate shock. These jumps, however, are largely a consequence of the oversimplified, two-point productivity process that we adopted for illustrative purposes, and are likely to disappear with a smoother productivity process. The forces combining into the observed dynamic path of wages are the following. First, the composition effect described in the previous paragraph

[11] Figure 3 also shows a slight dip in average labor productivity at the beginning of some of the expansions, immediately after the initial jump directly caused by the shock. This dip is also due to a composition effect: as argued before, at early stages of an expansion, when the unemployed are many, new hires get allocated disproportionately into low-productivity firms, which tends to bring down aggregate productivity. If λ_0 (which governs the inflow of unemployed job applicants) is sufficiently large, that effect may initially dominate the positive effect of labor upgrading on mean output per worker. But then precisely because λ_0 is large, the pool of unemployed job applicants becomes depleted very quickly, so that the effect described in this footnote is short-lived.

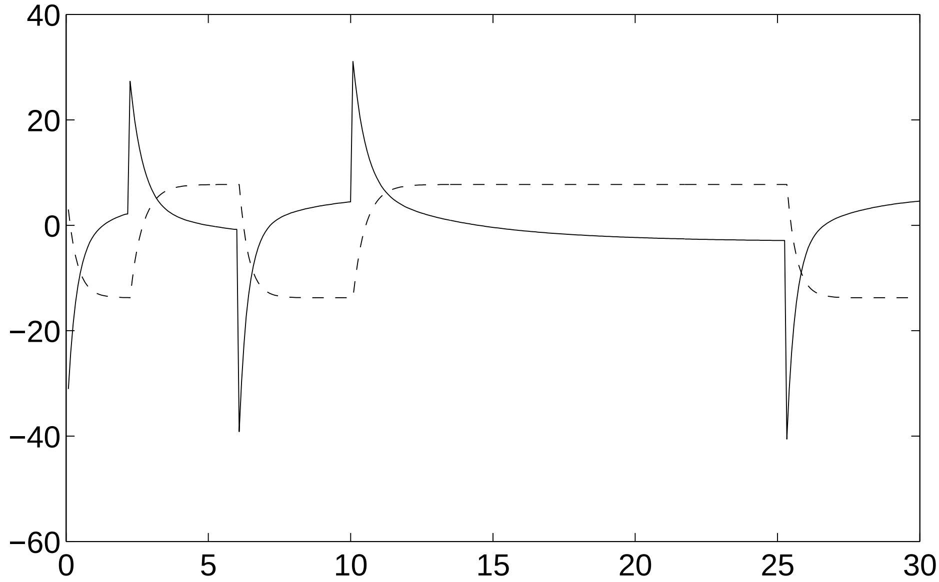

Fig. 4: Mean wage (in log, solid) and unemployment rate (dashed)

as governing the smooth dynamics of productivity also affects wages: more productive firms pay higher wages at all dates and workers move up the productivity (or wage) ladder in expansions and down that ladder in recessions. In addition, each firm-level wage also follows a dynamic path of its own. The fact mentioned in the discussion of Proposition 2 that larger firms have a motive to transfer higher values to their workers also applies across dates for a given firm: firms tend to post higher values whenever they grow in size, and vice versa.

7 Conclusion

This paper is the first to characterize stochastic equilibrium of an economy where the Law of One Price fails due to random search frictions and monopsony power, a problem that was long held to be intractable. Specifically, we introduce aggregate productivity shocks in a wage-posting model a la Burdett and Mortensen (1998), and we allow for rich state-contingent employment contracts. By extending the theory of Monotone Comparative Statics to a Dynamic Programming environment, we establish that, under mild sufficient conditions, the equilibrium is unique, constrained efficient, and very tractable, and we provide an algorithm to compute it. We show that the equilibrium stochastic dynamics of this model economy exhibit qualitative properties that are in line with the new business cycle facts illustrated in MPV09 and MPV10, most notably, that small firms as a group have relative less cyclical employment and returns to capital than large firms.

To conclude, we revisit some of our assumptions, both to explore the robustness of our theoretical results, which are the main focus of this paper, and to prepare the ground for future research on extensions of the model, with an eye also to improve its quantitative performance at business cycle frequencies.

We first observe that the unique RPE and constrained efficient allocation of employment is also the unique equilibrium allocation of more restricted contract-posting games. For example, suppose firms can post wages that are conditioned on aggregate productivity ω and own productivity θ, but not on own size L or employment distribution Λ because, say, L is too difficult to verify for the worker, and Λ is too difficult to measure for either the worker or the firm. Proposition 2 shows that a firm's best response to any strategy adopted by other firms is increasing in θ as long as it can be conditioned on the aggregate state ω. Therefore, this game in restricted contracts must also have a unique equilibrium which is RP. Constrained efficient reallocation can be implemented in relatively simple strategies that only depend on θ and ω.

The assumption of commitment to state-contingent wages is typically binding, as firms would like to renege on their promises once workers are hired. This assumption is standard in the literature, where it is well understood that commitment may both be beneficial to the firm and sustainable in long term employment relationships. Nonetheless, this assumption could be relaxed in many different ways. For example, the firm may be allowed a choice of whether or not commit to a specific strategy (as in Postel-Vinay and Robin, 2004). Or, it may only be committed to end-of-current-period payments with or (as in Coles, 2001) without an equal-treatment constraint. Then, a firm may choose to deliver value to its workers upon hiring them and then squeeze them against the participation constraint. Under the equal-treatment constraint, this strategy would not poach any workers. Without the constraint, this strategy of extremely front-loaded payments (pure sign-up bonus) has the cost of accelerating turnover. It is then plausible that an individual firm may not want to play this strategy if no other firm does, but rather choose to backload wages sequentially, even without commitment, due to the need to retain workers once hired. This is an interesting avenue for future research.

We have treated job-contact and job-destruction probabilities as exogenous, albeit state-dependent, objects. More natural and common is to endogenize them through a matching function and possibly idiosyncratic productivity shocks. On the first count, we envision the following natural extension of the model. A firm can post vacancies, or spend hiring effort, at a convex cost. Own vacancies determine the firm's sampling weight in workers' job search. The firm now has two tools at its disposal to recruit workers, promised contract value and vacancies (hiring effort). We expect the dynamic single-crossing property that we uncover in the value function of the optimal contract-posting problem to imply that not only the value offered to the worker, but also the intensity of hiring effort, increase with firm productivity and size. This result would only reinforce the mechanism that we highlight and which gives rise to the unique Rank-Preserving equilibrium. Quantitatively, it would greatly help to explain the empirical inequality in firm sizes based on labor turnover frictions alone. On the time domain, in order for large firms to exhibit more cyclical job creation rates in equilibrium, they would have to post in equilibrium a measure of vacancies that is procyclical relative to those of small firms.

We are aware that multiple factors, beyond just employment frictions, contribute to determine the size of a firm, most notably adjustment costs to other factors, such as capital adjustment costs and financial frictions, and diminishing marginal revenues from hiring, due to either technology (decreasing returns to labor due to another fixed factor), span of control frictions, or price-making power. Diminishing returns in wage-posting models have been partially explored in a steady state context, and can invalidate some of the equilibrium properties, such as the absence of atoms in the offer distribution. We cannot identify, though, obvious reasons why they would overturn the main result that equilibrium must be Rank-Preserving. It seems inconsistent with equilibrium logic that a more productive firm may decide to hire so many more workers as to drive its marginal revenue of labor strictly below that of a less productive firm, thus breaking the RP property.

We have treated firm productivity as a fixed, time-invariant parameter. Empirical evidence shows that revenue-based measures of firm productivity are subject to shocks, which determine entry, exit and a firm's life cycle. Shocks to firm productivity create obvious issues for Rank-Preserving equilibria, as a very large and productive firm may suddenly become unproductive, and then face contrasting incentives to offer its employees a high value, its sheer size and retention needs against low productivity. The relevant question, however, is how variable is a typical firm's productivity at business cycle frequencies. The scant empirical evidence in this regard, limited to the manufacturing sector, (Haltiwanger et al., 2008, Table 3) suggests that establishment-level TFP calculated from physical output data is very persistent, about as much as aggregate TFP. MPV09 show with data from several countries that several correlated features of a firm, such as its size, the average wages it pays, and its revenue-based productivity, when measured at one point in time strongly predict how job creation by the same firm responds to business cycle shocks that hit it *over two decades later*. This striking phenomenon suggests that our assumption of fixed firm productivity might be a reasonable approximation for our purposes.

Finally, our numerical example illustrates the practicality of these model economies as tools for business cycle analysis, and can also claim some success at matching old and new facts about aggregate labor markets, both cross-sectional and on the time domain. We believe that the slow propagation of aggregate shocks, as manifest in the sluggish response of average labor productivity and wages in our model, due to the slow upgrading of labor through job to job quits, is an important feature of actual business cycles which is missing altogether from existing quantitative business cycle models. The model extensions mentioned above, as well as possibly others, are bound to help to further improve the quantitative performance of the model. Overall, the results of this paper unlock an exciting research program.

References

[1] Barlevy, G. (2008) "Identification of Search Models using Record Statistics", *Review of Economic Studies*, 75(1) 29-64.

[2] Bontemps, C., J.-M. Robin, and G. J. Van den Berg (2000) "Equilibrium Search with Continuous Productivity Dispersion: Theory and Non-Parametric Estimation", *International Economic Review*, 41 (2) 305-58.

[3] Burdett, K. and M. Coles (2003) "Equilibrium Wage-Tenure Contracts", *Econometrica*, 71(5), 1377-1404.

[4] Burdett, K. and D. T. Mortensen (1998) "Wage Differentials, Employer Size and Unemployment", *International Economic Review*, 39, 257-73. [BM]

[5] Cahuc, P., F. Postel-Vinay and J.-M. Robin (2006) "Wage Bargaining with On-the-job Search: Theory and Evidence", *Econometrica*, 74(2), 323-64.

[6] Caputo, M. R. (2003) "The Comparative Dynamics of Closed-loop Controls for Discounted Infinite-Horizon Optimal Control Problems", *Journal of Economic Dynamics and Control*, 27, 1335-65.

[7] Carrillo-Tudela, C. (2009) "An Equilibrium Search Model when Firms Observe Workers' Employment Status", *International Economic Review*, 50(2), 485-506.

[8] Coles, M. G. (2001) "Equilibrium Wage Dispersion, Firm Size and Growth", *Review of Economic Dynamics*, 4, 159-87.

[9] Corbæ, D., M. B. Stinchcombe, and J. Zeman (2009), *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*, Princeton University Press.

[10] Dey, M. S. and C. J. Flinn (2005) "An Equilibrium Model of Health Insurance Provision and Wage Determination", *Econometrica*, 73(2), 571-627.

[11] Hall, R.E. (2005) "Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years", *in* M. Gertler and K. Rogoff, eds., *NBER Macroeconomics Annual*, 20, 1-137.

[12] Haltiwanger, J., L. Foster and C. Syverson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?", *American Economic Review*, 98:1, 394425.

[13] Menzio, G. and S. Shi (2008) "Efficient Search on the Job and the Business Cycle", PIER Working Paper 08-029.

[14] Mortensen, D. (2003) *Wage Dispersion: Why Are Similar Workers Paid Differently?*, MIT Press.

[15] Moscarini, G. and F. Postel-Vinay (2008) "The Timing of Labor Market Expansions: New Facts and a New Hypothesis", *in* D. Acemoglu, K. Rogoff and M. Woodford, eds., *NBER Macroeconomics Annual*, 23, 1-51. [MPV08]

[16] Moscarini, G. and F. Postel-Vinay (2009) "Large Employers Are More Cyclically Sensitive", NBER Working Paper 14740. [MPV09]

[17] Pissarides, C. (2000), *Equilibrium Unemployment Theory*, 2nd edition, MIT Press.

[18] Postel-Vinay, F. and J.-M. Robin (2002) "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70 (6), 2295-350.

[19] Postel-Vinay, F. and J.-M. Robin (2004) "To Match Or Not To Match? Optimal Wage Policy with Endogenous Worker Search Intensity", *Review of Economic Dynamics*, 7(2), 297-331.

[20] Robin, J.-M. (2009) "Labour Market Dynamics with Sequential Auctions: On Countercyclical Wage Inequality and Unemployment Volatility", Walras-Bowley Lecture, North American Summer Meeting of the Econometric Society, Boston University, June 2009.

[21] Royden, H. L. (1988) *Real Analysis – Third Edition*, Prentice-Hall.

[22] Rudanko, L. (2008) "Aggregate and Idiosyncratic Risk in a Frictional Labor Market", mimeo Boston University.

[23] Rudin, W. (1976) *Principles of Mathematical Analysis – Third Edition*, McGraw-Hill.

[24] Shimer, R. (2008) "Discussion of Moscarini and Postel-Vinay (2008)", *in* D. Acemoglu, K. Rogoff and M. Woodford, eds., *NBER Macroeconomics Annual*, 23, 59-71.

[25] Spear, S. and S. Srivastava, (1987), "On Repeated Moral Hazard With Discounting", *Review of Economic Studies*, 54, 599-617.

[26] Stevens, M. (2004) "Wage-Tenure Contracts in a Frictional Labour Market: Firms' Strategies for Recruitment and Retention", *Review of Economic Studies*, 71, 535-551.

[27] Topkis, D. M. (1998), *Supermodularity and Complementarity*, Princeton, N.J.: Princeton University Press.

Appendix

A Proof of Proposition 2

For convenience, we repeat the firm's DP problem (14):

$$
\begin{aligned}
S(\theta, L, \omega, \Lambda) &= \omega\theta L + \beta \int_\Omega \Bigg\{ \delta^{\omega'} U(\omega', \mathscr{T}(\omega', \Lambda)) L \\
&+ \sup_{W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))} \Bigg\langle S(\theta, \mathscr{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathscr{T}(\omega', \Lambda)) + L\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v \mid \omega', \Lambda) \\
&- W(\omega') \left(\lambda_0^{\omega'} \left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G(W(\omega') \mid \omega', \Lambda)\right) \Bigg\rangle \Bigg\} H(d\omega' \mid \omega),
\end{aligned}
$$

and the claim: if this problem has a solution, then any measurable selection $V(\theta, L, \omega, \Lambda)$ from the optimal correspondence is such that $V(\theta, L^*(\theta), \omega, \Lambda)$ is increasing in θ. We introduce the following notation:

$$
\begin{aligned}
A(\theta, L, \omega, \Lambda) &:= \omega\theta L + \beta \int_\Omega \delta^{\omega'} U(\omega', \mathscr{T}(\omega', \Lambda)) LH(d\omega' \mid \omega), \\
B(L, \omega', \Lambda; W(\omega')) &:= L\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v \mid \omega', \Lambda) \\
&\quad - W(\omega') \left(\lambda_0^{\omega'} \left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G(W(\omega') \mid \omega', \Lambda)\right).
\end{aligned}
$$

Our proof strategy is as follows. First, we define certain supermodularity properties SM of a value function that imply that the maximizer V in (14) is increasing in θ. Then, we fix an arbitrary Λ and show that the Bellman operator in (14) for the restricted problem with fixed Λ is a contraction mapping from the space of SM functions into itself, and that this space is Banach and closed under the sup norm. Therefore, for any fixed Λ (14) has a unique solution. Finally, *if* there exists a solution S to (14) when Λ is not fixed, then S must also solve the restricted problem (14) for any fixed Λ. By uniqueness and SM of the solution to the restricted problem any solution to the unrestricted problem must also have the SM properties. We cannot extend the same logic to show existence of S with variable Λ because Blackwell's sufficient conditions for a contraction mapping apply only to functions over $\mathbb{R}^n$.

So fix Λ to be some given CDF over $\left[\underline{\theta}, \overline{\theta}\right]$. Then, for any function $\mathscr{S}(\theta, L, \omega)$, we define the following operator $\mathbf{M}^{\mathbf{\Lambda}}$:

$$\begin{aligned}\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}(\theta, L, \omega) &:= A(\theta, L, \omega, \Lambda) \\ &+ \beta \int_{\Omega} \max_{W(\omega')} \Big\langle \mathscr{S}\left[\theta, \mathscr{L}(L, \omega', \Lambda, W(\omega')), \omega'\right] + B(L, \omega', \Lambda; W(\omega')) \Big\rangle H(d\omega' \mid \omega). \quad (27)\end{aligned}$$

The following additional consideration simplifies the proof: the worker participation constraint $W \geq U$ can be ignored in the proof. To see why, observe the following. Once we establish that an interior solution is increasing in θ, we can conclude that any set of firms that offers a corner solution $W = U$ and shuts down must be the set of the least productive firms. But then, the global solution, including the corner, is weakly increasing in θ as claimed. Incidentally, if all firms offered U, from the previous reasoning (and barring the trivial case where all firms are too unproductive to operate) the most productive firms would deviate and profitably offer more, so there exist always some firms that have an interior solution where PC does not bind.

Lemma 1 *Let $\mathscr{S}(\theta, L, \omega)$ be bounded, continuous in θ and L, increasing and convex in L and with increasing differences in (θ, L) over $\left(\underline{\theta}, \overline{\theta}\right) \times (0, 1)$. Then:*

1. $\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}$ *is bounded and continuous in θ and L;*
2. *There exists a measurable selection $V(\theta, L, \omega, \Lambda)$ from the maximizing correspondence associated with $\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}$;*
3. *Any such measurable selection V is increasing in θ and L;*
4. $\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}$ *is increasing and convex in L and with increasing differences in (θ, L) over $\left(\underline{\theta}, \overline{\theta}\right) \times (0, 1)$.*

Proof. In this proof, wherever possible without causing confusion, we will make the dependence of all functions on aggregate state variables ω and Λ implicit to streamline the notation.

Points 1 and 2 of this lemma are immediate: continuity of $\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}$ is a direct consequence of Berge's Theorem. Boundedness of $\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}$ is obvious by construction. Existence of a measurable selection from the maximizing correspondence associated with $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ is a direct consequence of the Measurable Selection Theorem.

To prove point 3, we first establish that the maximand in (27) has increasing differences in (θ, W) and (L, W). Monotonicity of V in θ and L will then follow from standard monotone comparative statics arguments. Proving that the maximand in (27) has increasing differences in (θ, W) is immediate as B is independent of θ: letting $\tau > 0$, differences in θ of the maximand equal $\mathscr{S}(\theta+\tau, \mathscr{L}(L, W)) - \mathscr{S}(\theta, \mathscr{L}(L, W))$ which is increasing in W because $\mathscr{L}$ is increasing in W by construction and $\mathscr{S}$ has increasing differences in (θ, L) by assumption. We thus now fix θ and focus on establishing that the maximand in (27) has increasing differences in (L, W). To this end, first note that, since $\mathscr{S}$ is assumed to be continuous and convex in L, it has left and right derivatives everywhere (and those two can at most differ at countably many points). Now take L and $h > 0$ and define the difference in L of the maximand in (27):

$$\begin{aligned}\mathscr{D}(W) := \mathscr{S}(\theta, \mathscr{L}(L+h, W)) - \mathscr{S}(\theta, \mathscr{L}(L, W)) \\ + h\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_W^{+\infty} v dF(v \mid \omega').\end{aligned}$$

(The dependence of $\mathscr{D}$ on θ is kept implicit.) We want to establish that $\mathscr{D}(W)$ is increasing in W. We do not know whether F and G, thus $\mathscr{D}$, are differentiable, so we proceed by showing that the upper-right Dini derivative of $\mathscr{D}(W)$, which we denote as $D^+\mathscr{D}(W)$ and which exists everywhere (although possibly equalling $\pm\infty$), is everywhere positive. Take $x > 0$:

$$\begin{aligned}&\mathscr{D}(W+x) - \mathscr{D}(W)\\
&\quad = \mathscr{S}(\theta, \mathscr{L}(L+h, W+x)) - \mathscr{S}(\theta, \mathscr{L}(L, W+x))\\
&\qquad\qquad - [\mathscr{S}(\theta, \mathscr{L}(L+h, W)) - \mathscr{S}(\theta, \mathscr{L}(L, W))]\\
&\qquad\qquad\qquad - h\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_W^{W+x} v dF(v \mid \omega')\\
&\quad = \{\mathscr{S}_{L,r}(\theta, \mathscr{L}(L+h, W)) + \varepsilon_1[\mathscr{L}(L+h, W+x) - \mathscr{L}(L+h, W)]\}\\
&\qquad\qquad\qquad \times \{\mathscr{L}(L+h, W+x) - \mathscr{L}(L+h, W)\}\\
&\quad - \{\mathscr{S}_{L,r}(\theta, \mathscr{L}(L, W)) + \varepsilon_2[\mathscr{L}(L, W+x) - \mathscr{L}(L, W)]\}\\
&\qquad\qquad \times \{\mathscr{L}(L, W+x) - \mathscr{L}(L, W)\}\\
&\qquad\qquad\qquad - h\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_W^{W+x} v dF(v \mid \omega')\end{aligned}$$

where ε_1 and ε_2 are functions that have limit 0 at 0, $f_{x,\ell}$ $[f_{x,r}]$ is used to designate the left [right] partial derivative of any function f w.r.t. x, and $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ has one-sided derivatives because it is convex. Majorizing the last integral:

$$\begin{aligned}
&\mathscr{D}(W+x)-\mathscr{D}(W)\\
&\quad\geq\left\{\mathscr{S}_{L,r}\left(\theta,\mathscr{L}(L+h,W)\right)+\varepsilon_1\left[\mathscr{L}(L+h,W+x)-\mathscr{L}(L+h,W)\right]\right\}\\
&\qquad\qquad\qquad\qquad\qquad\times\left\{\mathscr{L}(L+h,W+x)-\mathscr{L}(L+h,W)\right\}\\
&\quad-\left\{\mathscr{S}_{L,r}\left(\theta,\mathscr{L}(L,W)\right)+\varepsilon_2\left[\mathscr{L}(L,W+x)-\mathscr{L}(L,W)\right]\right\}\\
&\qquad\qquad\qquad\qquad\qquad\times\left\{\mathscr{L}(L,W+x)-\mathscr{L}(L,W)\right\}\\
&\qquad\qquad-h\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}(W+x)\left[F(W+x\mid\omega')-F(W\mid\omega')\right].
\end{aligned}$$

Dividing through by x and taking the limit superior as $x\to 0^+$ (using the definition of $\mathscr{L}$, the fact that $\mathscr{S}_{L,r}\geq 0$ by assumption, continuity of F and G, and some basic properties of Dini derivatives), we obtain:

$$\begin{aligned}
&D^+\mathscr{D}(W)\geq\\
&\mathscr{S}_{L,r}\left(\theta,\mathscr{L}(L+h,W)\right)\cdot\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{(L+h)\,D^+F(W)+\Lambda\left(\overline{\theta}\right)D^+G(W)\right\}\\
&\quad-\mathscr{S}_{L,r}\left(\theta,\mathscr{L}(L,W)\right)\cdot\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{LD_+F(W)+\Lambda\left(\overline{\theta}\right)D_+G(W)\right\}\\
&\qquad\qquad-h\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}WD_+F(W),
\end{aligned}$$

where, in standard fashion, D_+F denotes the lower-right Dini derivative of F (and likewise for G). Because F and G are increasing, their Dini derivatives are such that $D^+F\geq D_+F\geq 0$ (and likewise for G). Because $\mathscr{S}$ is convex in L by assumption, $\mathscr{S}_{L,r}$ is increasing in L. Combining all those properties, the latter inequality implies:

$$D^+\mathscr{D}(W)\geq\left[\mathscr{S}_{L,r}\left(\theta,\mathscr{L}(L,W)\right)-W\right]\cdot\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}hD_+F(W). \tag{28}$$

The only way the RHS in this last inequality can be negative is if $\mathscr{S}_{L,r}\left(\theta,\mathscr{L}(L,W)\right)-W<0$. We now show that this cannot be if W is an optimal selection. Let V be an optimal selection and let $x>0$. Optimality requires that:

$$\begin{aligned}
0\geq\;&\mathscr{S}\left(\theta,\mathscr{L}(L,V-x)\right)+L\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V-x}^{+\infty}vdF(v\mid\omega')\\
&\quad-(V-x)\left(\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)G(V-x\mid\omega')\right)\\
&\qquad-\mathscr{S}\left(\theta,\mathscr{L}(L,V)\right)-L\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V}^{+\infty}vdF(v\mid\omega')\\
&\qquad\qquad+V\left(\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)G(V\mid\omega')\right).
\end{aligned}$$

Collecting terms and again majorizing the integral term as we did for $\mathscr{D}$:

$$
\begin{aligned}
0 \geq \{\mathscr{S}_{L,\ell}(\theta, \mathscr{L}(L,V)) + \varepsilon\left[\mathscr{L}(L,V-x) - \mathscr{L}(L,V)\right]\}\cdot\{\mathscr{L}(L,V-x) - \mathscr{L}(L,V)\} \\
+ L\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}(V-x)\left[F(V-x \mid \omega') - F(V \mid \omega')\right] \\
- V\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)\left[G(V-x \mid \omega') - G(V \mid \omega')\right] \\
+ x\left(\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)G(V \mid \omega')\right).
\end{aligned}
$$

Now again taking the limit superior as $x \to 0^+$ (in what follows D^-F and D_-F designate the upper and lower left Dini derivative of F, respectively, and likewise for G):[12]

$$
\begin{aligned}
0 \geq -\mathscr{S}_{L,\ell}(\theta, \mathscr{L}(L,V)) \cdot \lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{LD_-F(V) + \Lambda\left(\overline{\theta}\right)D_-G(V)\right\} \\
+ V\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{LD^-F(V) + \Lambda\left(\overline{\theta}\right)D^-G(V)\right\} \\
+ \lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)G(V \mid \omega').
\end{aligned}
$$

Finally recalling that $D^-F \geq D_-F \geq 0$ (and likewise for G), the latter inequality implies:

$$
\mathscr{S}_{L,\ell}(\theta, \mathscr{L}(L,V)) - V \geq \frac{\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)G(V \mid \omega')}{\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{LD_-F(V) + \Lambda\left(\overline{\theta}\right)D_-G(V)\right\}} \geq 0. \tag{29}
$$

This, together with (28), shows that $D^+\mathscr{D}(V) \geq 0$ at all V which is an optimal selection, i.e. at all V in the support of F. To finally establish that $\mathscr{D}$ is increasing over the support of F, recall that, as F and G are continuous by Proposition 1, so is $W \mapsto \mathscr{L}(L,W)$. Moreover, as $\mathscr{S}$ is convex in L (by assumption), it is continuous w.r.t. L. Thus by inspection, $\mathscr{D}$ is a continuous function of W. Continuity plus the fact that $D^+\mathscr{D}(V) \geq 0$ are sufficient to ensure that $\mathscr{D}$ is strictly increasing (see, e.g., Proposition 2 p99 in Royden, 1988). Point 3 of the lemma is thus proven.

Now on to point 4. Take $(\theta_0, L_0) \in \left(\underline{\theta}, \overline{\theta}\right) \times (0,1)$ and $h > 0$ such that $(\theta_0 + h, L_0 + h)$ are still in $\left(\underline{\theta}, \overline{\theta}\right) \times (0,1)$. We first consider right-differentiability of $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ w.r.t. L at L_0. Again fixing an arbitrary selection V from the optimal policy correspondence, we note that, while V may have a discontinuity at (θ_0, L_0), the fact that it is increasing in L ensures that $V\left(\theta_0, L_0^+, \omega'\right) :=$

[12]This uses the facts that $\mathscr{S}_{L,\ell} \geq 0$, that F and G are continuous, and that $D^-[-f] = -D^-f$ for any function f.

$\lim_{h\to 0^+} V(\theta_0, L_0+h, \omega')$ exists everywhere (and likewise for $V\left(\theta_0^+, L_0, \omega'\right)$). By point 3, $V\left(\theta_0, L_0^+, \omega'\right)$ is increasing in L_0. Then:

$$
\begin{aligned}
&\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}(\theta_0, L_0+h) - \mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^+\right) = A(\theta_0, L_0+h) - A(\theta_0, L_0) \\
&\quad + \beta \int_\Omega \Big\langle \mathscr{S}\left[\theta_0, \mathscr{L}(L_0+h, V(\theta_0, L_0+h, \omega'))\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^+, \omega'\right)\right)\right] \\
&\qquad + B(L_0+h; V(\theta_0, L_0+h, \omega')) - B\left(L_0; V\left(\theta_0, L_0^+, \omega'\right)\right)\Big\rangle H(d\omega' \mid \omega) \\
&\geq A(\theta_0, L_0+h) - A(\theta_0, L_0) \\
&\quad + \beta \int_\Omega \Big\langle \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0+h, V\left(\theta_0, L_0^+, \omega'\right)\right)\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^+, \omega'\right)\right)\right] \\
&\qquad + B\left(L_0+h; V\left(\theta_0, L_0^+, \omega'\right)\right) - B\left(L_0; V\left(\theta_0, L_0^+, \omega'\right)\right)\Big\rangle H(d\omega' \mid \omega) \\
&= \left(\omega\theta_0 + \beta \int_\Omega \delta^{\omega'} U(\omega') H(d\omega' \mid \omega)\right) \cdot h \\
&\quad + \beta \int_\Omega \Bigg\langle \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0+h, V\left(\theta_0, L_0^+, \omega'\right)\right)\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^+, \omega'\right)\right)\right] \\
&\qquad\qquad + h \cdot \left(1-\delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V\left(\theta_0, L_0^+, \omega'\right)}^{+\infty} v\, dF(v \mid \omega') \Bigg\rangle H(d\omega' \mid \omega), \qquad (30)
\end{aligned}
$$

where the last equality follows from the definitions of A and B. Then again:

$$
\begin{aligned}
&\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}(\theta_0, L_0+h) - \mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^+\right) = A(\theta_0, L_0+h) - A(\theta_0, L_0) \\
&\quad + \beta \int_\Omega \Big\langle \mathscr{S}\left[\theta_0, \mathscr{L}(L_0+h, V(\theta_0, L_0+h, \omega'))\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^+, \omega'\right)\right)\right] \\
&\qquad + B(L_0+h; V(\theta_0, L_0+h, \omega')) - B\left(L_0; V\left(\theta_0, L_0^+, \omega'\right)\right)\Big\rangle H(d\omega' \mid \omega) \\
&\leq A(\theta_0, L_0+h) - A(\theta_0, L_0) \\
&\quad + \beta \int_\Omega \Big\langle \mathscr{S}\left[\theta_0, \mathscr{L}(L_0+h, V(\theta_0, L_0+h, \omega'))\right] - \mathscr{S}\left[\theta_0, \mathscr{L}(L_0, V(\theta_0, L_0+h, \omega'))\right] \\
&\qquad + B(L_0+h; V(\theta_0, L_0+h, \omega')) - B(L_0; V(\theta_0, L_0+h, \omega'))\Big\rangle H(d\omega' \mid \omega) \\
&= \left(\omega\theta_0 + \beta \int_\Omega \delta^{\omega'} U(\omega') H(d\omega' \mid \omega)\right) \cdot h \\
&\quad + \beta \int_\Omega \Bigg\langle \mathscr{S}\left[\theta_0, \mathscr{L}(L_0+h, V(\theta_0, L_0+h, \omega'))\right] - \mathscr{S}\left[\theta_0, \mathscr{L}(L_0, V(\theta_0, L_0+h, \omega'))\right] \\
&\qquad\qquad + h \cdot \left(1-\delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V(\theta_0, L_0+h, \omega')}^{+\infty} v\, dF(v \mid \omega') \Bigg\rangle H(d\omega' \mid \omega). \qquad (31)
\end{aligned}
$$

Now dividing through by h in (30) and (31), and invoking continuity w.r.t. V of $\mathscr{L}_L(L, V) = \left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{F}(V)\right)$ (by continuity of F), everywhere right-differentiability of $\mathscr{S}$ w.r.t. L (by convexity of $\mathscr{S}$), and existence of a right limit of V at any L_0 (by monotonicity of V established in point 1 of this lemma), we see that the lower and upper bounds of $\frac{1}{h}\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}(\theta_0, L_0+h) - \mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^+\right)\right]$ exhibited in (30) and (31) both converge to the same limit as $h \to 0^+$, which, together with continuity of $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ in L at L_0 which implies $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^+\right) = \mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}(\theta_0, L_0)$, establishes right-differentiability of $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ w.r.t L with the following expression for $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}(\theta, L)$

$$\begin{aligned}\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}(\theta, L) &= \omega\theta + \beta\int_{\Omega}\delta^{\omega'}U(\omega')\,H(d\omega' \mid \omega)\\&\quad + \beta\int_{\Omega}\Big\langle \mathscr{S}_{L,r}\left[\theta, \mathscr{L}\left(L, V\left(\theta, L^+, \omega'\right)\right)\right]\cdot\mathscr{L}_L\left(L, V\left(\theta, L^+, \omega'\right)\right)\\&\qquad + \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V(\theta,L^+,\omega')}^{+\infty} v\,dF(v \mid \omega')\Big\rangle H(d\omega' \mid \omega).\end{aligned} \tag{32}$$

Straightforward inspection shows that $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}(\theta, L) > 0$, so that $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ is increasing in L. We now show that $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}(\theta, L)$ is increasing in L and θ. It is sufficient to show that the term under the $\int$ in (32) is increasing in L and θ for all $\omega' \in \Omega$. We begin with L. Let $L_1 < L_2 \in [0,1]^2$. To lighten the notation, let $V_k = V\left(\theta, L_k^+, \omega'\right)$ for $k = 1, 2$. Because V is increasing in L, $V_2 \geq V_1$. Then:

$$\begin{aligned}&\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right]\cdot\mathscr{L}_L(L_2, V_2) - \mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_1, V_1)\right]\cdot\mathscr{L}_L(L_1, V_1)\\&\qquad\qquad - \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V_1}^{V_2} v\,dF(v \mid \omega')\\&\quad = \left[\mathscr{L}_L(L_2, V_2) - \mathscr{L}_L(L_1, V_1)\right]\cdot\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right]\\&\qquad + \mathscr{L}_L(L_1, V_1)\cdot\left(\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right] - \mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_1, V_1)\right]\right)\\&\qquad\qquad - \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V_1}^{V_2} v\,dF(v \mid \omega')\\&\quad = \mathscr{L}_L(L_1, V_1)\cdot\left(\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right] - \mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_1, V_1)\right]\right)\\&\qquad\qquad + \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V_1}^{V_2}\left(\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right] - v\right)dF(v \mid \omega'),\end{aligned}$$

where the last equality stems from the definition of $\mathscr{L}_L$. Because $\mathscr{S}_{L,r}$ and $\mathscr{L}$ are both increasing in L, and because $\mathscr{L}$ is also increasing in V, the first term in the r.h.s. of the last equality above is positive. Finally, convexity of $\mathscr{S}$ combined with the first-order condition (29) implies that $\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right] \geq \mathscr{S}_{L,\ell}\left[\theta, \mathscr{L}(L_2, V_2)\right] \geq V_2$, so that $\mathscr{S}_{L,r}\left[\theta, \mathscr{L}(L_2, V_2)\right] \geq v$ for all $v \leq V_2$, implying that the integral term is nonnegative. This shows that $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}$ is (strictly) increasing in L. The proof that $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}$ is strictly increasing in θ proceeds along similar lines (details available upon request). Thus $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ is a continuous function whose right partial derivative w.r.t. L exists everywhere, is increasing in L — which proves convexity w.r.t. L —, and increasing in θ — which proves increasing differences in (θ, L). $\square$

Now consider the set of functions defined over $\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega$ that are continuous in (θ, L) and call it $C_{\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega}$. That set is a Banach space when endowed with the sup norm. As Lemma 1 suggests we will be interested in the properties a subset $C'_{\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega} \subset C_{\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega}$ of functions that are increasing and convex in L and have increasing differences in (θ, L). We next prove two ancillary lemmas, which will establish as a corollary (Corollary 1) that $C'_{\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega}$ is closed in $C_{\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega}$ under the sup norm.[13]

Lemma 2 *Let X be an interval in $\mathbb{R}$ and $f_n : X \to \mathbb{R}$, $N \in \mathbb{N}$ such that $\{f_n\}$ converges uniformly to f. Then:*

1. *if f_n is nondecreasing for all n, so is f;*
2. *if f_n is convex for all n, so is f.*

Proof. For point 1, take $(x_1, x_2) \in X^2$ such that $x_2 > x_1$. Fix $k \in \mathbb{N}$. By uniform convergence, $\exists n_k \in \mathbb{N} : \forall n \geq n_k$, $\forall x \in X$, $|f_n(x) - f(x)| < \frac{1}{2k}$. Then:

$$f(x_2) - f(x_1) = \underbrace{f(x_2) - f_{n_k}(x_2)}_{>-1/2k} + \underbrace{f_{n_k}(x_2) - f_{n_k}(x_1)}_{\geq 0 \text{ by monotonicity of } f_{n_k}} + \underbrace{f_{n_k}(x_1) - f(x_1)}_{>-1/2k} > -\frac{1}{k}.$$

As the above is valid for an arbitrary choice of $k \in \mathbb{N}$ and $(x_1, x_2) \in X^2$, it establishes that f is nondecreasing. For point 2, uniform convergence of $\{f_n\}$ to f implies pointwise convergence, so that Theorem 6.2.35 p282 in Corbae, Stinchcombe and Zeman (2009) can be applied. $\square$

Lemma 3 *Let $X \subset \mathbb{R}^2$ be a convex set and $f_n : X \to \mathbb{R}$, $N \in \mathbb{N}$ be functions with increasing differences such that $\{f_n\}$ converges uniformly to f. Then f has increasing differences.*

Proof. Let $\{(x_1, y_1), (x_2, y_2)\} \in X^2$ such that $x_2 > x_1$ and $y_2 > y_1$. Fix $k \in \mathbb{N}$. By uniform convergence, $\exists n_k \in \mathbb{N} : \forall n \geq n_k$, $\forall (x, y) \in X$, $|f_n(x, y) - f(x, y)| <$

[13] While for the purposes of this proof (which is concerned with closedness under the sup norm) both lemmas are stated for sequences that converge uniformly, it is straightforward to extend them to the case of pointwise convergent sequences.

$\frac{1}{4k}$. Then:

$$
\begin{aligned}
&f(x_2,y_2)-f(x_1,y_2)\\
&=\underbrace{f(x_2,y_2)-f_{n_k}(x_2,y_2)}_{>-1/4k}+\underbrace{f_{n_k}(x_2,y_2)-f_{n_k}(x_1,y_2)}_{>f_{n_k}(x_2,y_1)-f_{n_k}(x_1,y_1)\text{ by ID of }f_{n_k}}+\underbrace{f_{n_k}(x_1,y_2)-f(x_1,y_2)}_{>-1/4k}\\
&>-\frac{1}{2k}+f_{n_k}(x_2,y_1)-f_{n_k}(x_1,y_1)\\
&=-\frac{1}{2k}+\underbrace{f_{n_k}(x_2,y_1)-f(x_2,y_1)}_{>-1/4k}+f(x_2,y_1)-f(x_1,y_1)+\underbrace{f(x_1,y_1)-f_{n_k}(x_1,y_1)}_{>-1/4k}\\
&>-\frac{1}{k}+f(x_2,y_1)-f(x_1,y_1).
\end{aligned}
$$

As the above is valid for an arbitrary choice of $k \in \mathbb{N}$ and $\{(x_1,y_1),(x_2,y_2)\} \in X^2$, it establishes that f has increasing differences. □

Corollary 1 *The set $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ of functions defined over $[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega$ that are increasing and convex in L and have increasing differences in (θ,L) is a closed subset of $C_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ under the sup norm.*

The latter corollary establishes that, given a fixed Λ, the set of functions that are relevant to Lemma 1 is a closed subset of a Banach space of functions under the sup norm. The following lemma shows that the operator considered in Lemma 1 is a contraction under that same norm.

Lemma 4 *The operator $\mathbf{M}^{\Lambda}$ defined in (27) maps $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ into itself and is a contraction of modulus β under the sup norm.*

Proof. That $\mathbf{M}^{\Lambda}$ maps $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ into itself flows directly from a subset of the proof of Lemma 1. To prove that $\mathbf{M}$ is a contraction, it is straightforward to check using (27) that $\mathbf{M}^{\Lambda}$ satisfies Blackwell's sufficient conditions with modulus β. □

We are now in a position to prove the proposition. Given the initially fixed Λ, the operator $\mathbf{M}^{\Lambda}$, which by Lemma 4 is a contraction from $C_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ into itself, and has a unique fixed point $\mathscr{S}_{\Lambda}$ in that set (by the Contraction Mapping Theorem). Moreover, since $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ is a closed subset of $C_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ (Lemma 2) and since $\mathbf{M}_{\Lambda}$ also maps $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ into itself (Lemma 1), that fixed point $\mathscr{S}_{\Lambda}$ belongs to $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$.

Summing up, what we have established thus far is that for any fixed $\Lambda \in C_{[\underline{\theta},\overline{\theta}]}$, the operator $\mathbf{M}^{\Lambda}$ over functions of (θ,L,ω) has a unique, bounded and continuous fixed point $\mathscr{S}^{\star}_{\Lambda} = \mathbf{M}_{\Lambda}\mathscr{S}^{\star}_{\Lambda} \in C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega} \subset C_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$.

We finally turn to the Bellman operator $\mathbf{M}$ which is relevant to the firm's problem. That operator $\mathbf{M}$ applies to functions $\overline{\mathscr{S}}$ defined on $\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega\times C_{\left[\underline{\theta},\overline{\theta}\right]}$ and is defined as the following "extension" of $\mathbf{M}^{\Lambda}$:

$$\mathbf{M}\overline{\mathscr{S}}\left(\theta,L,\omega,\Lambda\right):=A\left(\theta,L,\omega,\Lambda\right)$$
$$+\beta\int_{\Omega}\max_{W(\omega')}\left\langle\overline{\mathscr{S}}\left[\theta,\mathscr{L}\left(L,\omega',\Lambda,W\left(\omega'\right)\right),\omega',\mathscr{T}\left(\omega,\Lambda\right)\right]+B\left(L,\omega',\Lambda;W\left(\omega'\right)\right)\right\rangle H\left(d\omega'\mid\omega\right).$$

If an equilibrium exists, then a firm has a best response and a value S which solves $S=\mathbf{M}S$. For every $\Lambda\in C_{\left[\underline{\theta},\overline{\theta}\right]}$, by definition of $\mathbf{M}$ and $\mathbf{M}_{\Lambda}$ this implies $S=\mathbf{M}_{\Lambda}S$. Since the fixed point of $\mathbf{M}_{\Lambda}$ is unique, if $S=\mathbf{M}S$ exists then for every fixed $\Lambda\in C_{\left[\underline{\theta},\overline{\theta}\right]}$ we have for all $(\theta,L,\omega)\in\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega$: $S\left(\theta,L,\omega,\Lambda\right)=\mathscr{S}^{\star}_{\Lambda}\left(\theta,L,\omega\right)$. Therefore, if the value function S and an equilibrium of the contract-posting game exist, then $S\in C'_{\left[\underline{\theta},\overline{\theta}\right]\times[0,1]\times\Omega}$: the typical firm's value function is continuous in θ and L, increasing and convex in L and has increasing differences in (θ,L). By the same standard monotone comparative statics arguments that we invoked in the proof of Lemma 1, the maximizing correspondence is increasing in θ and L in the strong set sense, hence all of its measurable selections are weakly increasing in θ and L.

The proposition is finally established by the following simple induction. Consider two firms with $\theta_2>\theta_1$. By assumption, at date 0, $L_2\geq L_1$. Because any selection $V\left(\theta,L,\omega,\Lambda\right)$ from the maximizing correspondence of the typical firm's problem is increasing in θ and L, the values posted by those two firms at date 0 are such that $V_2\geq V_1$. Then because $\mathscr{L}$ is strictly increasing in both L and V, firm 2 is again larger than firm 1 at date 1. The same reasoning applies again at date 1, and at all subsequent dates, so that $V_2\geq V_1$ holds true at all dates.

Finally, if firms are equally productive the RP property follows as a simple corollary of the convexity of S in L, by the assumption that the initial Λ is continuous. □

We conclude with a remark on atoms in the initial size distribution, including the symmetric case of identical firms that are equally productive and start out with the same size. That case would require mixed strategies in the first period. After the mixing plays out, in the second period of play firms would differ in size, and the previous case would apply from then on. We leave the computation of the equilibrium mixed strategies to future research.

B Proof of Proposition 3

In an attempt to simplify the notation without causing confusion, we define:

$$V^{\star}\left(\theta,\omega\right):=V\left(\theta,L^{\star}\left(\theta\right),\omega,\Lambda^{\star}\right)$$

for use throughout this proof. This notation keeps the dependence of $V(\cdot)$ on Λ implicit.

The main purpose of Proposition 3 is actually to establish claim 2, continuous differentiability of $V^{\star}$. Our proof strategy is as follows. We know from Proposition 2 that the optimal policy $V^{\star}$ is increasing in θ, hence differentiable a.e. It remains to show that it is differentiable everywhere. To do so, first, we establish continuity properties of $V(\theta, L, \omega, \Lambda^{\star})$ in θ, both for fixed L and for $L = L^{\star}(\theta)$, and in L at $L = L^{\star}(\theta)$ for fixed θ. Using these properties, we show that any solution to the Bellman equation when all other firms are playing a RPE, $S(\theta, L, \omega, \Lambda^{\star})$, is continuously differentiable in L at $L = L^{\star}(\theta)$; that is, on the equilibrium path the shadow marginal value of one worker always exists and is continuous in firm size. Next, we exploit this property and the implications of RPE to show that the optimal policy $V^{\star}$ is Lipschitz continuous in θ. This implies that $V^{\star}$ is differentiable everywhere.

We begin with an ancillary lemma, which is interesting in its own right.

Lemma 5 *V has the following continuity properties along the (RP) Equilibrium path:*

1. $\theta \mapsto V(\theta, L^{\star}(\theta), \omega, \Lambda^{\star}) = V^{\star}(\theta, \omega)$ *is continuous;*

2. $L \mapsto V(\theta, L, \omega, \Lambda^{\star})$ *is continuous at* $L = L^{\star}(\theta)$*;*

3. $\tau \mapsto V(\tau, L^{\star}(\theta), \omega, \Lambda^{\star})$ *is continuous at* $\tau = \theta$.

Proof. $\theta \mapsto V^{\star}(\theta, \omega)$ is increasing by Proposition 2, so $V^{\star}$ can only have (countably many) jump discontinuities. But then a jump discontinuity in $V^{\star}$ would imply a gap in the support of F, which is inconsistent with equilibrium as argued in Appendix A. This proves claim 1 of the lemma.

For claim 2, fix θ and $\varepsilon > 0$. Then by continuity of $V^{\star}$ (point 1 of this lemma), $\exists\, \alpha > 0 : \forall \eta \in (0, \alpha]$, $V^{\star}(\theta, \omega) \leq V^{\star}(\theta + \eta, \omega) \leq V^{\star}(\theta, \omega) + \varepsilon$. But then monotonicity of V in L and in θ (see Appendix A) further implies: $V^{\star}(\theta, \omega) \leq V(\theta, L^{\star}(\theta + \eta), \omega, \Lambda^{\star}) \leq V^{\star}(\theta + \eta, \omega) \leq V^{\star}(\theta, \omega) + \varepsilon$, so that $\forall\, L \in [L^{\star}(\theta), L^{\star}(\theta + \eta)]$, $V(\theta, L, \omega, \Lambda^{\star}) - V(\theta, L^{\star}(\theta), \omega, \Lambda^{\star}) \leq \varepsilon$, which establishes right-continuity of V in L at $L^{\star}(\theta)$. Left-continuity is established in the same way, and so is continuity of $\tau \mapsto V(\tau, L^{\star}(\theta), \omega, \Lambda^{\star})$ at $\tau = \theta$. $\square$

We now go on to establish point 1 of the proposition. In so doing, to avoid notational overload, we will keep the dependence of all value functions and laws of motion on $\Lambda^{\star}$ implicit. Now first, convexity of S w.r.t. L was established as a by-product of Proposition 2 (see Appendix A), and implies that S is everywhere left- and right-differentiable w.r.t. L, and that the right and left derivatives $S_{L,r}$ and $S_{L,\ell}$ are both increasing functions of L. As such they have right and left limits everywhere. We can thus define $S_{L,r}(\theta, L^{+}, \omega) =$

$\lim_{h\to 0+} S_{L,r}(\theta, L+h, \omega)$, and symmetrically $S_{L,\ell}(\theta, L^-, \omega) = \lim_{h\to 0+} S_{L,\ell}(\theta, L-h, \omega)$. Now following exactly the same steps as in (30) and (31) (see the proof of Lemma 1 in Appendix A), only applied to S, we establish:

$$
\begin{aligned}
S_{L,r}\left(\theta, L^+, \omega\right) = \omega\theta + \beta \int_\Omega \delta^{\omega'} U(\omega') H(d\omega' \mid \omega) \\
+ \beta \int_\Omega \Big\langle S_{L,r}\left[\theta, \mathscr{L}\left(L, \omega', V\left(\theta, L^+, \omega'\right)\right), \omega'\right] \cdot \mathscr{L}_L\left(L, \omega', V\left(\theta, L^+, \omega'\right)\right) \\
+ \left(1-\delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V(\theta, L^+, \omega')}^{+\infty} v dF(v \mid \omega') \Big\rangle H(d\omega' \mid \omega).
\end{aligned}
$$

Next, the facts that V is increasing in L (see the proof of Proposition 2) and continuous in L at $L = L^\star(\theta)$ (from Lemma 5), combined with continuity of $\mathscr{L}$ and $\mathscr{L}_L$ w.r.t. V (by continuity of F), imply that $\mathscr{L}_L(L, \omega', V(\theta, L^+, \omega')) = \mathscr{L}_L(L, \omega', V(\theta, L, \omega'))$ and $S_{L,r}[\theta, \mathscr{L}(L, \omega', V(\theta, L^+, \omega')), \omega'] = S_{L,r}\left[\theta, \mathscr{L}(L, \omega', V(\theta, L, \omega'))^+, \omega'\right]$ at $L = L^\star(\theta)$. As a further consequence:

$$
\begin{aligned}
S_{L,r}\left(\theta, L^\star(\theta)^+, \omega\right) = \omega\theta + \beta \int_\Omega \delta^{\omega'} U(\omega') H(d\omega' \mid \omega) \\
+ \beta \int_\Omega \Big\langle S_{L,r}\left[\theta, \mathscr{L}\left(L, \omega', V^\star(\theta, \omega')\right)^+, \omega'\right] \cdot \mathscr{L}_L\left(L, \omega', V^\star(\theta, \omega')\right) \\
+ \left(1-\delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V^\star(\theta, \omega')}^{+\infty} v dF(v \mid \omega') \Big\rangle H(d\omega' \mid \omega). \quad (33)
\end{aligned}
$$

A symmetric expression can be arrived at in the same way for $S_{L,\ell}\left(\theta, L^\star(\theta)^-, \omega\right)$, so that defining $\mathscr{D}_{S_L}(\theta, L, \omega) := S_{L,r}(\theta, L^+, \omega) - S_{L,\ell}(\theta, L^-, \omega)$, which is positive by convexity of S in L, we have:

$$
\begin{aligned}
0 \leq \mathscr{D}_{S_L}(\theta, L^\star(\theta), \omega) &= \beta \int_\Omega \mathscr{D}_{S_L}\left[\theta, \mathscr{L}\left(L, \omega', V^\star(\theta, \omega')\right), \omega'\right] \cdot \mathscr{L}_L\left(L, \omega', V^\star(\theta, \omega')\right) H(d\omega' \mid \omega) \\
&< \beta \int_\Omega \mathscr{D}_{S_L}\left[\theta, \mathscr{L}\left(L, \omega', V^\star(\theta, \omega')\right), \omega'\right] H(d\omega' \mid \omega).
\end{aligned}
$$

At this point, if we can prove that $\mathscr{D}_{S_L}$ is uniformly bounded above by some $K > 0$, then iterating the last inequality will show that $0 \leq \mathscr{D}_{S_L}(\theta, L^\star(\theta), \omega) < \beta^n K$ for all $n \in \mathbb{N}$, which implies that $\mathscr{D}_{S_L}(\theta, L^\star(\theta), \omega) = 0$ for all (θ, ω) and that S_L exists everywhere. Since S is convex, S_L is increasing, hence it can only have jumps up. But we just concluded that its right and left limit are equal everywhere, so S_L is continuous for all $L \in \left[0, \bar{L}\right)$, thus proving point 1 of the proposition.

We still need to show that $\mathscr{D}_{S_L}$ is uniformly bounded above. Because $S_{L,\ell} \geq 0$, it suffices to show that $S_{L,r}$ is bounded above. The following series of inequalities use the facts that

$S_{L,r}$ is increasing in L (by convexity of S), that $L \leq 1$ (since the total mass of workers in the economy is 1), and that $S_{L,r}(\theta, L^\star(\theta), \omega) \geq S_{L,\ell}(\theta, L^\star(\theta), \omega) \geq V^\star(\theta, \omega) \geq U(\omega)$ again invoking convexity in conjunction with the FOC (29):

$$
\begin{aligned}
&S_{L,r}\left(\theta, L^+, \omega\right) \leq S_{L,r}(\theta, 1, \omega) \\
&\quad \leq \omega\theta + \beta \int_\Omega \delta^{\omega'} U(\omega') H(d\omega' \mid \omega) + \beta \int_\Omega \left\langle S_{L,r}(\theta, 1, \omega') \cdot \left(1 - \delta^{\omega'}\right)\left(1 - \lambda_1^{\omega'} \overline{F}(V^\star(\theta, \omega'))\right) \right. \\
&\qquad \left. + \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V^\star(\theta,\omega')}^{+\infty} v dF(v \mid \omega') \right\rangle H(d\omega' \mid \omega) \\
&\quad \leq \omega\theta + \beta \int_\Omega \delta^{\omega'} U(\omega') H(d\omega' \mid \omega) + \beta \int_\Omega \left(1 - \delta^{\omega'}\right) S_{L,r}(\theta, 1, \omega') H(d\omega' \mid \omega) \\
&\quad \leq \omega\theta + \beta \int_\Omega S_{L,r}(\theta, 1, \omega') H(d\omega' \mid \omega).
\end{aligned}
$$

This establishes that $S_{L,r}(\theta, 1, \omega) \leq \max_\Omega \omega\theta / (1 - \beta)$ for all ω, and so $\max_\Omega \omega\theta / (1 - \beta)$ is also a uniform upper bound for $S_{L,r}(\theta, L^\star(\theta), \omega)$. This completes the proof of point 1 in the proposition.

We now go straight to point 3 before proving point 2. Consider the problem of a firm choosing W to best-respond to all other firms playing a RPE. By a simple improvement argument, $W \in \left[V^\star(\underline{\theta}, \omega'), V^\star\left(\overline{\theta}, \omega'\right)\right]$. Since $V^\star$ is continuous and increasing, then offering any such best response W is equivalent to choosing a type τ to imitate such that $W = V^\star(\tau, \omega')$. In any RPE, by Proposition 2, the best response by a firm θ of current size $L^\star(\theta)$ is 'truthful revelation', $\tau^\star = \theta$, which solves

$$
\begin{aligned}
S(\theta, L^\star(\theta), \omega) &= A(\theta, L^\star(\theta), \omega) \\
&\quad + \beta \int_\Omega \max_{\tau(\omega')} \left\langle S\left[\theta, \mathscr{L}(L^\star(\theta), \omega', \tau(\omega')), \omega'\right] + B(L^\star(\theta), \omega', \tau(\omega')) \right\rangle H(d\omega' \mid \omega)
\end{aligned}
$$

where, with a slight abuse of notation:

$$
\begin{aligned}
\mathscr{L}(L, \omega', \tau) = L\left(1 - \delta^{\omega'}\right)&\left(1 - \lambda_1^{\omega'} \overline{F}(V^\star(\tau, \omega') \mid \omega')\right) \\
&+ \lambda_0^{\omega'}\left(1 - \Lambda^\star\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right) \Lambda^\star\left(\overline{\theta}\right) G(V^\star(\tau, \omega') \mid \omega')
\end{aligned}
$$

and

$$
\begin{aligned}
B(L, \omega', \tau) = L&\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V^\star(\tau,\omega')}^{+\infty} v dF(v \mid \omega') \\
&- V^\star(\tau, \omega')\left(\lambda_0^{\omega'}\left(1 - \Lambda^\star\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right) \Lambda^\star\left(\overline{\theta}\right) G(V^\star(\tau, \omega') \mid \omega')\right)
\end{aligned}
$$

are continuous functions of L and τ. Using the RP property

$$\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\tau\right)=L^{\star}\left(\theta\right)\left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{\Gamma}\left(\tau\right)\right)+\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}\left(\tau\right) \tag{34}$$

$$\begin{aligned}B\left(L^{\star}\left(\theta\right),\omega',\tau\right)=&L^{\star}\left(\theta\right)\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{\tau}^{\overline{\theta}}V^{\star}\left(x,\omega'\right)d\Gamma\left(x\right)\\&-V^{\star}\left(\tau,\omega'\right)\left(\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}\left(\tau\right)\right).\end{aligned}$$

Lemma 6 *$V^{\star}$ is Lipschitz continuous, hence absolutely continuous and* $V^{\star}\left(\theta,\omega'\right)=\int^{\theta}V^{\star\prime}\left(x,\omega'\right)dx$.

Proof. Fix θ and ω'. Optimality requires for all $h>0$:

$$\begin{aligned}&S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta-h,\omega'\right)\right]+L^{\star}\left(\theta\right)\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{\theta-h}^{\overline{\theta}}V^{\star}\left(x,\omega'\right)d\Gamma\left(x\right)\\&\qquad-V^{\star}\left(\theta-h,\omega'\right)\left(\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}\left(\theta-h\right)\right)\\&\quad\le S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta\right),\omega'\right]+L^{\star}\left(\theta\right)\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{\theta}^{\overline{\theta}}V^{\star}\left(x,\omega'\right)d\Gamma\left(x\right)\\&\qquad\qquad-V^{\star}\left(\theta,\omega'\right)\left(\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}\left(\theta\right)\right).\end{aligned}$$

Rearranging:

$$\begin{aligned}&\left[V^{\star}\left(\theta,\omega'\right)-V^{\star}\left(\theta-h,\omega'\right)\right]\cdot\left[\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}\left(\theta\right)\right]\\&\le S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta\right),\omega'\right]-S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta-h\right),\omega'\right]\\&-\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}\left(\theta\right)\int_{\theta-h}^{\theta}V^{\star}\left(x,\omega'\right)d\Gamma\left(x\right)+V^{\star}\left(\theta-h,\omega'\right)\cdot\left[\Lambda^{\star}\left(\theta\right)-\Lambda^{\star}\left(\theta-h\right)\right]\right\}\\&\le S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta\right),\omega'\right]-S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta-h\right),\omega'\right]\\&-V^{\star}\left(\theta-h,\omega'\right)\cdot\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}\left(\theta\right)\left[\Gamma\left(\theta\right)-\Gamma\left(\theta-h\right)\right]+\left[\Lambda^{\star}\left(\theta\right)-\Lambda^{\star}\left(\theta-h\right)\right],\right\}\end{aligned} \tag{35}$$

where the second inequality is obtained by minorizing the integral term, remarking that $V^{\star}$ is increasing in θ. Now differentiability of S w.r.t. L and the definition (34) together imply that:[14]

$$\begin{aligned}S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\theta,\omega'\right),\omega'\right]-S\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\theta-h,\omega'\right),\omega'\right]=S_L\left[\theta,\mathscr{L}\left(L^{\star}\left(\theta\right),\theta,\omega'\right),\omega'\right]\\\times\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}\left(\theta\right)\left[\Gamma\left(\theta\right)-\Gamma\left(\theta-h\right)\right]+\left[\Lambda^{\star}\left(\theta\right)-\Lambda^{\star}\left(\theta-h\right)\right]\right\}+o\left(h\right).\end{aligned}$$

[14]The $o(h)$ term at the end comes from $o\left(\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta\right)-\mathscr{L}\left(L^{\star}\left(\theta\right),\omega',\theta-h\right)\right)=o\left(L^{\star}\left(\theta\right)\left[\Gamma\left(\theta\right)-\Gamma\left(\theta-h\right)\right]+\left[\Lambda^{\star}\left(\theta\right)-\Lambda^{\star}\left(\theta-h\right)\right]\right)=o\left(h\right)$ by differentiability of Γ and $\Lambda^{\star}$.

Substituting into (35), we obtain:

$$\begin{aligned}
&\left[V^{\star}(\theta,\omega')-V^{\star}(\theta-h,\omega')\right]\cdot\left[\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)\right]\\
&\qquad\le\left\{S_L\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\omega',\theta\right),\omega'\right]-V^{\star}(\theta-h,\omega')\right\}\\
&\quad\times\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}(\theta)\left[\Gamma(\theta)-\Gamma(\theta-h)\right]+\left[\Lambda^{\star}(\theta)-\Lambda^{\star}(\theta-h)\right]\right\}+o(h)
\end{aligned}$$

Now dividing through by $h>0$ and taking the limit superior:

$$0\le D^{-}V^{\star}(\theta)\le 2\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)L^{\star}(\theta)\,\gamma(\theta)\,\frac{S_L\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\omega',\theta\right),\omega'\right]-V^{\star}(\theta,\omega')}{\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)}$$

(the first inequality is a direct consequence of $V^{\star}$ being increasing). All terms in the r.h.s. are uniformly bounded above ($\gamma(\theta)$ by assumption and S_L by a property established earlier in this proof). So $V^{\star}$ is a continuous function with bounded (upper-left) Dini derivative, which is sufficient to ensure Lipschitz-continuity (see, e.g., Problem 20.c p112 in Royden, 1988). □

We are finally in a position to prove claim 2 of the proposition, namely that in any RPE, $V^{\star}(\theta)$ is continuously differentiable. Because $V^{\star}$ is increasing, we know that $V_{\theta}^{\star}$ exists outside of a null set, say N_V. Therefore, for each $\theta\in\left[\underline{\theta},\overline{\theta}\right]\backslash N_V$ we can take a derivative in the Bellman equation and write a NFOC:

$$V_{\theta}^{\star}(\theta,\omega')=2\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)L^{\star}(\theta)\,\gamma(\theta)\,\frac{S_L\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\omega',\theta\right),\omega'\right]-V^{\star}(\theta,\omega')}{\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)}:=q(\theta)\,.$$

Notice that the RHS $q(\theta)$ is continuous in θ, where it exists, i.e. in the set $\left[\underline{\theta},\overline{\theta}\right]\backslash N_V$ (recall that $L^{\star}$ is continuous by the assumption that L_0 is). Since $\left[\underline{\theta},\overline{\theta}\right]\backslash N_V$ is the complement of a set of measure 0, it is dense in $\left[\underline{\theta},\overline{\theta}\right]$. Therefore, for all $\theta\in N_V$, there exists a sequence $\{\theta_n\}$, $\theta_n\in\left[\underline{\theta},\overline{\theta}\right]\backslash N_V$ such that $\theta_n\to\theta$. As $V_{\theta}^{\star}(\theta_n,\omega')$ exists and equals $q(\theta_n)$ for all θ_n in this sequence, using the NFOC and continuity of q: $\lim_{n\to\infty}V_{\theta}^{\star}(\theta_n,\omega')=\lim_{n\to\infty}q(\theta_n)=q(\theta)$. Let

$$\tilde{V}_{\theta}(\theta):=\begin{cases}V_{\theta}^{\star}(\theta,\omega') & \theta\in\left[\underline{\theta},\overline{\theta}\right]\backslash N_V\\ q(\theta) & \text{otherwise}\end{cases}$$

which, by the last argument, is continuous everywhere. Then,

$$V^{\star}(\theta,\omega')=V^{\star}(\underline{\theta},\omega')+\int_{\underline{\theta}}^{\theta}V_{\theta}^{\star}(x,\omega')\,dx=V^{\star}(\underline{\theta},\omega')+\int_{\underline{\theta}}^{\theta}\tilde{V}_{\theta}(x)\,dx$$

where the second equality follows from the fact that $V_{\theta}^{\star}(\theta,\omega')\neq\tilde{V}_{\theta}(\theta)$ only on a null set. So $V^{\star}(\theta,\omega')$ is the integral of a continuous function $\tilde{V}_{\theta}$, hence it is differentiable with $V_{\theta}^{\star}(\theta,\omega')=\tilde{V}_{\theta}(\theta)$ everywhere, and the FOC $V_{\theta}^{\star}(\theta,\omega')=q(\theta)$ holds everywhere. Point 3 of the proposition is thus proven.

Finally on to point 2. Introducing a time index τ and using the notation

$$\begin{aligned}
\Lambda^\star_{\tau+1} &= \mathscr{T}\left(\omega_{\tau+1},\Lambda^\star_\tau\right)\\
\Delta\left(\theta,\omega_\tau\right) &= \left(1-\delta^{\omega_\tau}\right)\left(1-\lambda_1^{\omega_\tau}\overline{\Gamma}\left(\theta\right)\right)\\
\mathscr{U}\left(\theta,\omega_\tau,\Lambda^\star_\tau\right) &= \mathbf{E}_{\omega_{\tau+1}|\omega_\tau}\Bigg[\delta^{\omega_{\tau+1}}U\left(\omega_{\tau+1},\Lambda^\star_{\tau+1}\right)\\
&\quad+\left(1-\delta^{\omega_{\tau+1}}\right)\lambda_1^{\omega_{\tau+1}}\int_\theta^{+\infty}V^\star\left(x\mid\omega_{\tau+1},\Lambda^\star_{\tau+1}\right)d\Gamma\left(x\right)\Bigg],\\
\text{and: }\ \mu\left(\theta,\omega_\tau,\Lambda^\star_\tau\right) &= S_L\left(\theta,L^\star_\tau\left(\theta\right),\omega_\tau,\Lambda^\star_\tau\right),
\end{aligned}$$

we can rewrite the Euler equation (33) as follows:

$$\mu\left(\theta,\omega_t,\Lambda^\star_t\right)=\omega_t\theta+\beta\mathscr{U}\left(\theta,\omega_t,\Lambda^\star_t\right)+\beta\mathbf{E}_{\omega_{t+1}|\omega_t}\left[\Delta\left(\theta,\omega_{t+1}\right)\mu\left(\theta,\omega_{t+1},\Lambda^\star_{t+1}\right)\right]. \tag{36}$$

For any measurable function ϕ of ω and any t, define recursively the linear operator $X^t_0[\phi]=\phi$ and

$$X^t_s[\phi]=\mathbf{E}_{\omega_{t+s}|\omega_{t+s-1}}\left[\Delta\left(\theta,\omega_{t+s}\right)X^t_{s-1}[\phi]\right]\ \text{ for } s=1,2\cdots n-1.$$

After n forward substitutions, we can write (36) as

$$\mu\left(\theta,\omega_t,\Lambda^\star_t\right)=\mu_n\left(\theta,\omega_t,\Lambda^\star_t\right)+\beta^{n+1}X^t_{n+1}[\mu]$$

where

$$\mu_n\left(\theta,\omega_t,\Lambda^\star_t\right)=\sum_{s=0}^n\beta^s\left\{X^t_{s+1}[\omega\theta]+\beta X^t_{s+1}\left[\mathscr{U}\left(\theta,\omega,\Lambda^\star\right)\right]\right\}.$$

Since $\mu>0$ and $\Delta\in(0,1)$ with probability 1

$$\begin{aligned}
0<&\mu\left(\theta,\omega_t,\Lambda^\star_t\right)-\mu_n\left(\theta,\omega_t,\Lambda^\star_t\right)=\left|\mu\left(\theta,\omega_t,\Lambda^\star_t\right)-\mu_n\left(\theta,\omega_t,\Lambda^\star_t\right)\right|=\beta^{n+1}X^t_{n+1}[\mu]\\
\leq&\beta^{n+1}\mathbf{E}_{\omega_{t+1}|\omega_t}\left[\mathbf{E}_{\omega_{t+2}|\omega_{t+1}}\left[\cdots\mathbf{E}_{\omega_{t+n}|\omega_{t+n-1}}\left[\mu\left(\theta,\omega_{t+n},\Lambda^\star_{t+n}\right)\right]\right]\right]\\
=&\ \beta^{n+1}\mathbf{E}_{\omega_{t+n}|\omega_t}\left[\mu\left(\theta,\omega_{t+n},\Lambda^\star_{t+n}\right)\right]
\end{aligned}$$

Since a firm can always guarantee itself positive profits and employment by offering its workers the value of unemployment, then $L^\star\left(\theta,\omega,\Lambda^\star\right)$ is bounded away from 0 with probability one. So the TVC (22) implies that, as $n\to\infty$, the last term vanishes, thus μ_n converges pointwise (and indeed uniformly) to μ.

Next, taking derivatives

$$\begin{aligned}
\frac{\partial\mu_n\left(\theta,\omega_t,\Lambda^\star_t\right)}{\partial\theta}&=\omega_t+\frac{\lambda_1^\omega\gamma\left(\theta\right)}{1-\lambda_1^\omega\overline{\Gamma}\left(\theta\right)}+\sum_{s=1}^n\beta^sX^t_s\left[\omega+\frac{\lambda_1^\omega\gamma\left(\theta\right)}{1-\lambda_1^\omega\overline{\Gamma}\left(\theta\right)}\omega\theta\right]\\
&\quad+\sum_{s=2}^{n+1}\beta^sX^t_s\left[\frac{\lambda_1^\omega\gamma\left(\theta\right)}{1-\lambda_1^\omega\overline{\Gamma}\left(\theta\right)}\mathscr{U}\left(\theta,\omega,\Lambda^\star\right)-\left(1-\delta^\omega\right)\lambda_1^\omega V^\star\left(\theta\mid\omega,\Lambda^\star\right)\gamma\left(\theta\right)\right]
\end{aligned}$$

which is continuous in θ. As the arguments of the operator X_s^t in the last expression are continuous in ω, θ on the compact set $\Omega \times \left[\underline{\theta}, \overline{\theta}\right]$, the $X_s^t[\cdot]$ terms in the sums are bounded above and below uniformly with probability 1 by some upper bound $X' < \infty$ and lower bound $-X'$ for all (θ, ω_t). Therefore, driven by discounting, the two sums converge as $n \to \infty$, and

$$\lim_{n\to\infty} \frac{\partial \mu_n(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta} = \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta}$$

exists for all $\theta, \omega_t, \Lambda_t^\star$. Also

$$\begin{aligned}\frac{\partial \mu_n(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta} - \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta} &= -\sum_{s=n+1}^{\infty} \beta^s X_s^t \left[\omega + \frac{\lambda_1^\omega \gamma(\theta)}{1-\lambda_1^\omega \overline{\Gamma}(\theta)} \omega\theta\right] \\ &- \sum_{s=n+2}^{\infty} \beta^s X_s^t \left[\frac{\lambda_1^\omega \gamma(\theta)}{1-\lambda_1^\omega \overline{\Gamma}(\theta)} \mathscr{U}(\theta, \omega, \Lambda^\star) - (1-\delta^\omega)\lambda_1^\omega V^\star(\theta \mid \omega, \Lambda^\star)\gamma(\theta)\right]\end{aligned}$$

So, for all $\theta, \omega_t, \Lambda_t^\star$, $|X_s^t[\cdot]| < X'$ implies

$$\left|\frac{\partial \mu_n(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta} - \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta}\right| < X' \frac{\beta^{n+1} + \beta^{n+2}}{1-\beta}$$

so that convergence of derivatives is uniform. By Theorem 7.17 in Rudin (1976), conclude that μ is continuously differentiable in θ with:

$$\frac{\partial \mu(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta} = \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t^\star)}{\partial \theta},$$

which completes the proof of the proposition. □

C Proof of Proposition 4

We first show that the limit in (24) exists, is positive and uniformly bounded above. By assumpion, $\omega\theta$ and b^ω are positive and uniformly bounded above by some $K < \infty$, therefore by the definition of $\mathbf{T}$, $\mathbf{T}\binom{\omega\theta}{b^\omega}$ is also positive and uniformly bounded above by K, and by induction the same is true of $\mathbf{T}^j\binom{\omega\theta}{b^\omega}$. Hence the sequence $\sum_{j=0}^{n-1} \beta^j \mathbf{T}^j \binom{\omega\theta}{b^\omega}$ is increasing and uniformly bounded above by $K/(1-\beta)$, so each of the two sums in this sequence must converge and the limit exists and is positive and bounded above by $K/(1-\beta)$.

We next show that, if there exists a RPE, then it is given by (24). Suppose there exists a RPE $\binom{\mu}{U}$. Then by definition of RPE $\mu \geq \mathbf{T}_V[\mu, U] \geq U \geq 0$ and $\mathbf{T}_V[\mu, U]$ is increasing in θ. Then, by inspection of $\mathbf{T}_\mu$: $\mu_t\left(\overline{\theta} \mid \omega\right) = \omega\overline{\theta} + \beta \mathbf{T}_\mu \binom{\mu}{U} \left(\overline{\theta} \mid \omega\right) \leq \omega\overline{\theta} + \beta \mathbf{E}_{\omega_{t+1}|\omega_t}\left[\mu_{t+1}\left(\overline{\theta} \mid \omega_{t+1}\right)\right]$. Since $\mathbf{E}_{\omega_{t+1}|\omega_t}[\mu_{t+1}] \geq 0$, iterating $\mathbf{T}$ forward and using the TVC:

$$0 \leq \beta^n \mathbf{T}_\mu^n[\mu, U]\left(\overline{\theta} \mid \omega\right) \leq \beta^n \mathbf{E}_{\omega_{t+n}|\omega_t}\left[\mu_{t+n}\left(\overline{\theta} \mid \omega_{t+n}\right)\right] \longrightarrow 0,$$

so that $\mu\left(\overline{\theta} \mid \omega\right) \leq \omega\overline{\theta}/\left(1-\beta\right) \leq K/\left(1-\beta\right)$. By Proposition 3, in any RPE $\partial\mu/\partial\theta$ exists; the proof of Proposition 2 shows that S has increasing differences in (θ, L), so that $\mu = S_L$ is increasing in θ. So for all θ, $0 \leq U \leq \mu\left(\theta \mid \omega\right) \leq K/\left(1-\beta\right)$ and $\binom{\mu}{U}$ is uniformly bounded above. By definition of a RPE, $\binom{\mu}{U}$ must solve (23). Substituting forward in (23) and using $U \leq \mu\left(\theta \mid \omega\right) \leq K/\left(1-\beta\right)$ we find $0 \leq \beta^n \mathbf{T}^n \binom{\mu}{U} \leq \beta^n K/\left(1-\beta\right) \to 0$, so that $\binom{\mu}{U} = \binom{\omega\theta}{b^\omega} + \beta\mathbf{T}\binom{\omega\theta}{b^\omega} + \beta^2\mathbf{T}^2\binom{\mu}{U} = \cdots = \binom{\mu^\star}{U^\star}$ as claimed.

We finally turn to existence. By construction, $\binom{\mu^\star}{U^\star}$ solves (23). Moreover, since $\binom{\mu^\star}{U^\star}$ is uniformly bounded above, it satisfies the TVC. So we only have to show that it satisfies $0 \leq U^\star \leq \mathbf{T}_V\left[\mu^\star, U^\star\right]$ and $\mathbf{T}_V\left[\mu^\star, U^\star\right]$ increasing in θ. By definition of $\mathbf{T}_V$:

$$\begin{aligned}\frac{\partial \mathbf{T}_V\left[\mu^\star, U^\star\right]}{\partial\theta}\left(\theta \mid \omega\right) &= \frac{\frac{\partial Q_t}{\partial\theta}\left(\theta \mid \omega\right)}{Q_t\left(\theta \mid \omega\right)}\left(\mu^\star\left(\theta \mid \omega\right) - \mathbf{T}_V\left[\mu^\star, U^\star\right]\left(\theta \mid \omega\right)\right) \\ &= \frac{\frac{\partial Q_t}{\partial\theta}\left(\theta \mid \omega\right)}{Q_t\left(\theta \mid \omega\right)}\left(\mu^\star\left(\theta \mid \omega\right) - U^\star\left(\omega\right) - \int_{\underline{\theta}}^{\theta}\left(\mu^\star\left(x \mid \omega\right) - U^\star\left(\omega\right)\right)\frac{\frac{\partial Q_t}{\partial\theta}\left(x \mid \omega\right)}{Q_t\left(\theta \mid \omega\right)}dx\right)\end{aligned}$$

so it suffices to prove that $\mu^\star\left(\underline{\theta} \mid \omega\right) \geq U\left(\omega\right)$ and $\mu^\star$ is increasing in θ. Now consider the sequence of functions $\left\{\mu_n, U_n\right\}_{n\in\mathbb{N}}$ defined by $\mu_0\left(\theta \mid \omega\right) = \omega\theta$, $U_0\left(\omega\right) = b^\omega$, and $\binom{\mu_{n+1}}{U_{n+1}} = \binom{\omega\theta}{b^\omega} + \beta\mathbf{T}\binom{\mu_n}{U_n}$. Clearly this sequence converges to $\binom{\mu^\star}{U^\star}$. Suppose μ_n is increasing in θ and greater than U_n for some n. It is straightforward to see from the definition of $\mathbf{T}_\mu$ that μ_{n+1} is increasing in θ. Then:

$$\mu_{n+1}\left(\theta \mid \omega\right) - U_{n+1}\left(\omega\right) = \left(\omega\theta - b^\omega\right) + \left(\mathbf{T}_\mu\left[\mu_n, U_n\right]\left(\theta \mid \omega\right) - \mathbf{T}_U\left[\mu_n, U_n\right]\left(\omega\right)\right).$$

The condition $\omega\underline{\theta} \geq b^\omega$ in all states ensures that the first term is positive, while the conditions $\left(1-\delta^\omega\right)\lambda_1^\omega \geq \lambda_0^\omega$ guarantees that the second terms is also positive. $\square$

D Details of the simulation algorithm

This appendix illustrates in detail the projection method that we use to solve for the shadow values μ and π in the PDE system (25,26) in each aggregate state $\omega_t = \omega$. Rewrite (25) as:

$$\begin{aligned}\frac{\partial\mu_t}{\partial\theta}\left(\theta \mid \omega\right) &= \omega + \beta\left(1-\delta^\omega\right)\left[\lambda_1^\omega\gamma\left(\theta\right)\pi_{t+1}\left(\theta \mid \omega\right) + \frac{\partial\mu_{t+1}}{\partial\theta}\left(\theta \mid \omega\right)\left(1 - \lambda_1^\omega\overline{\Gamma}\left(\theta\right)\right)\right] \\ &\quad + \beta\sigma^\omega\left\{\left(1-\delta^{\omega'}\right)\left[\lambda_1^{\omega'}\gamma\left(\theta\right)\pi_{t+1}\left(\theta \mid \omega'\right) + \frac{\partial\mu_{t+1}}{\partial\theta}\left(\theta \mid \omega'\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}\left(\theta\right)\right)\right]\right. \\ &\quad \left. - \left(1-\delta^\omega\right)\left[\lambda_1^\omega\gamma\left(\theta\right)\pi_{t+1}\left(\theta \mid \omega\right) + \frac{\partial\mu_{t+1}}{\partial\theta}\left(\theta \mid \omega\right)\left(1 - \lambda_1^\omega\overline{\Gamma}\left(\theta\right)\right)\right]\right\}. \end{aligned} \quad (37)$$

As explained in the main text, the idea is to approximate cross-state jumps in those shadow values by a known function. The "jump term" that we need to approximate is the last term in curly brackets in (37), that we denote by $J(\theta, L^{\star}(\theta), \omega, \Lambda^{\star}(\theta) \mid \mathbf{a})$. Given a specific vector of coefficients $\mathbf{a}$, system (26,37), together with the transversality condition (22), becomes a pair of independent systems of PDEs, one for each aggregate state, which can be separately numerically solved over the infinite future for any initial value of $(\Lambda^{\star}(\cdot), L^{\star}(\cdot))$ using the algorithm described in MPV08. We thus proceed in the following steps:

0. Pick an initial state of the economy $(\omega_0, \Lambda_0(\cdot), L_0(\cdot))$ and simulate a path of ω. Denote switching dates as $(s_1, s_2, \dots)$.

1. Fix a parameter $\mathbf{a}$.

2. Given the choice of $\mathbf{a}$ made at step 1 and the implied J-function, solve (26,37,22) using the appropriate initial conditions. More specifically:

   (a) Solve (26,37,22) with initial condition $(\Lambda_0(\cdot), L_0(\cdot))$ as if state ω_0 prevailed forever. This implies certain values for $\partial\mu_0(\theta \mid \omega_0)/\partial\theta$, $\partial\mu_{s_1}(\theta \mid \omega_0)/\partial\theta$, $\pi_0(\theta \mid \omega_0)$, $\pi_{s_1}(\theta \mid \omega_0)$ and $\left(\Lambda^{\star}_{s_1}(\cdot), L^{\star}_{s_1}(\cdot)\right)$ at date s_1 when the next aggregate shock occurs.

   (b) Solve (26,37,22) with initial condition $\left(\Lambda^{\star}_{s_1}(\cdot), L^{\star}_{s_1}(\cdot)\right)$ as if state ω_1 prevailed over $t \in [s_1, +\infty)$. This implies certain values for $\partial\mu_{s_1}(\theta \mid \omega_1)/\partial\theta$, $\partial\mu_{s_2}(\theta \mid \omega_1)/\partial\theta$, $\pi_{s_1}(\theta \mid \omega_1)$, $\pi_{s_2}(\theta \mid \omega_1)$ and $\left(\Lambda^{\star}_{s_2}(\cdot), L^{\star}_{s_2}(\cdot)\right)$.

   (c) Solve (26,37,22) with initial condition $\left(\Lambda^{\star}_{s_2}(\cdot), L^{\star}_{s_2}(\cdot)\right)$ as if state ω_2 prevailed over $t \in [s_2, +\infty)$, etc. That is, repeat step 2, mutatis mutandis, for the first K jumps in ω (in practice with a two-state process for ω, two jumps — one up, one down — are enough).

3. The simulations in step 2 provide a vector of jumps in $\partial\mu/\partial\theta$: for $k = 1, \cdots, K$

$$\left(1-\delta^{\omega'}\right)\left[\lambda_1^{\omega'}\gamma(\theta)\,\pi_{s_k}(\theta \mid \omega') + \frac{\partial\mu_{s_k}}{\partial\theta}(\theta \mid \omega')\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\right]$$
$$- (1-\delta^{\omega})\left[\lambda_1^{\omega}\gamma(\theta)\,\pi_{s_k}(\theta \mid \omega) + \frac{\partial\mu_{s_k}}{\partial\theta}(\theta \mid \omega)\left(1-\lambda_1^{\omega}\overline{\Gamma}(\theta)\right)\right]. \quad (38)$$

   Compare those with the jumps predicted from the initially chosen function $J(\cdot \mid \mathbf{a})$ and the simulated path of $(\omega, L^{\star}(\cdot), \Lambda^{\star}(\cdot))$. If different, update $\mathbf{a}$ and start over at step 1. Exactly how $\mathbf{a}$ is updated depends on the chosen functional form for the approximate jump function J. In practice we use a projection on polynomials,[15] and the updated vector of coefficients $\mathbf{a}$ is obtained by regression of the "simulated jumps" in (38) on the elements of J.

[15] After many trials, a good compromise between accuracy and speed of convergence was found using projection on degree-five polynomials in $\theta \times \Lambda^{\star}(\theta)$ and $\theta \times L^{\star}(\theta)$. With that specification the root mean squared prediction error is in the order of 1/100th of a percent of the mean (absolute value) simulated jump.